

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号  
特開2002-57981  
(P2002-57981A)

(43)公開日 平成14年2月22日(2002.2.22)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	テームト*(参考)
H 0 4 N 5/91		G 1 1 B 27/00	A 5 C 0 1 8
G 1 1 B 27/00		H 0 4 N 5/91	N 5 C 0 5 3
H 0 4 N 5/7826		5/782	A 5 D 1 1 0
5/92		5/92	H

審査請求 未請求 請求項の数28 O L 外国語出願 (全 66 頁)

(21)出願番号 特願2001-115802(P2001-115802)  
(22)出願日 平成13年4月13日(2001.4.13)  
(31)優先権主張番号 5 4 8 6 8 5  
(32)優先日 平成12年4月13日(2000.4.13)  
(33)優先権主張国 米国 (U S)

(71)出願人 000005496  
富士ゼロックス株式会社  
東京都港区赤坂二丁目17番22号  
(72)発明者 リン ディー、ウィルコックス  
アメリカ合衆国 94308 カリフォルニア  
州 バロ アルト スタンフォード アベ  
ニュー 555  
(72)発明者 パトリック ディーウ  
アメリカ合衆国 94025 カリフォルニア  
州 メンロ パーク ユニバーシティ ド  
ライブ 564 ナンバー3  
(74)代理人 100079049  
弁理士 中島 淳 (外1名)

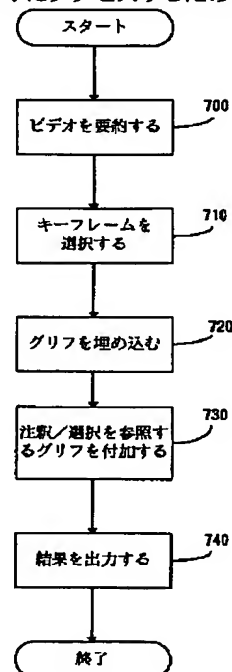
最終頁に続く

(54)【発明の名称】 データストリームにアクセスするインタフェース、データストリームへのアクセス用検索の形成方法、データストリームアクセス方法及びノートからビデオにアクセスするためのデバイス

(57)【要約】

【課題】 ノート作成中に監視されるビデオ又は他のデータストリームの重要部分を単純且つ正確な検索メカニズムで提供する。

【解決手段】 録画したビデオは、ビデオから導出されたプリントしたノート又はサマリーからアクセスされる。録画したビデオを分析することによってサマリーが自動的に形成され、デジタルインク及びビデオでノートを作成するデバイスのユーザによって注釈が付けられる。ノート及びサマリーは、録画したビデオへの時間ベースの索引又はオフセットを提供するデータグリフと共にプリントされる。索引又はオフセットは、プリントアウトのグリフを走査することによって検索される。グリフ情報は多くの方法でプリントアウトに埋め込まれることができる。アクセスしたビデオはノート作成デバイス又はウェブブラウザタイプのデバイスのサマリーインタフェースで再生される。



## 【特許請求の範囲】

【請求項1】 データストリームの少なくとも一部分に対応する少なくとも一つのオブジェクトと、前記オブジェクトに対応する前記データストリームの一部分を特定する各オブジェクトに関連するコードと、を含む出力、を含む、データストリームにアクセスするインタフェース。

【請求項2】 前記コードが対応する各オブジェクトに埋め込まれる請求項1に記載のインタフェース。

【請求項3】 前記出力が、前記コードの各々を前記データストリームのオフセットに関連させる表をさらに含む請求項1に記載のインタフェース。

【請求項4】 前記出力がプリントアウトである請求項1に記載のインタフェース。

【請求項5】 前記データストリームがビデオである請求項1に記載のインタフェース。

【請求項6】 前記コードがデータグリフである請求項1に記載のインタフェース。

【請求項7】 前記表がデータグリフである請求項3に記載のインタフェース。

【請求項8】 前記コードの各々がデータグリフカーベットを含み、前記オブジェクトは、前記データグリフカーベットの対応するコードに対応する位置に配置される、請求項1に記載のインタフェース。

【請求項9】 前記データグリフカーベットのコードが、前記データストリームの適切なオフセットも表すタイムライン上のポイントに対応するように、前記データグリフカーベットが前記タイムラインに関連して配置される請求項8に記載のインタフェース。

【請求項10】 データストリームに関連する少なくとも一つのオブジェクトを特定するステップと、関連する各オブジェクトに対応する前記データストリームのオフセットを決定するステップと、各オフセット及び対応する関連オブジェクトを特定する、索引を形成するステップと、を有するデータストリームへのアクセス用索引の形成方法。

【請求項11】 対応する前記オフセットを特定するコードに関連する特定した前記オブジェクトの各々を出力するステップをさらに含む請求項10に記載の方法。

【請求項12】 前記データストリームが、ビデオ又はマルチメディアデータファイルである請求項10に記載の方法。

【請求項13】 前記オブジェクトが、前記データストリームのキーフレーム、前記データストリームの選択したフレーム及び前記データストリームを参照して付けられたユーザ注釈のうちの少なくとも一つである請求項10に記載の方法。

【請求項14】 関連するデジタルインクストロークの

セットをユーザ注釈としてグループ化するステップをさらに含む請求項13に記載の方法。

【請求項15】 前記グループ化するステップが、デジタルインクストロークのセットへのグループ化をさらに含み、任意の一つのグループの各デジタルインクストロークが、所定の時間間隔内でユーザによって形成される、請求項13に記載の方法。

【請求項16】 データストリームに対応するオブジェクトを走査するステップと、走査した前記オブジェクトに基づいて前記データストリームのオフセットを決定するステップと、決定した前記オフセットに対応する前記データストリームの一部分にアクセスするステップと、を有するデータストリームアクセス方法。

【請求項17】 前記決定するステップが、走査した前記オブジェクトに関連するコードを読み取るステップと、関連する前記コードをオフセットに関連させる表から該オフセットにアクセスするステップと、を含む請求項16に記載の方法。

【請求項18】 前記アクセスするステップが、前記コードに対応する前記データストリームの一部分に対する要求を遠隔コンピュータに送出するステップと、前記データストリームの前記部分を受信するステップと、を含む請求項16に記載の方法。

【請求項19】 前記要求が、http、ftp又は他のインターネット関連要求タイプのうちの一つを含み、前記遠隔コンピュータがインターネットに接続され、少なくとも前記データストリームの少なくとも前記部分へのアクセスを有する、請求項18に記載の方法。

【請求項20】 前記要求が、前記オブジェクトに関連して読み取られた前記コードを含み、前記遠隔コンピュータが、前記コードを前記データストリームの対応する部分に関連させる表にアクセスする、請求項18に記載の方法。

【請求項21】 前記受信ステップが、ストリーミングビデオの受信を含む請求項16に記載の方法。

【請求項22】 ビデオ索引に埋め込まれたグリフコードを走査するように構成された走査デバイスと、走査した前記グリフコードを読み取り、該グリフコードに基づいて前記ビデオの一部分にアクセスするように構成された演算デバイスと、検索した前記ビデオの部分を出力するように構成されたディスプレイメカニズムと、を含むノートからビデオにアクセスするためのデバイス。

【請求項23】 前記演算デバイスが、ワイヤード又はワイヤレスネットワークの何れかに接続できるネットワーク接続を含み、前記演算メカニズムが、前記ビデオの前記部分を遠隔デ

バイスから前記ネットワーク接続を介して要求できるように構成されている、

請求項22に記載のデバイス。

【請求項24】 前記演算メカニズムが、前記グラフィックコードを前記ビデオのオフセットに相関させる表を含む請求項23に記載のデバイス。

【請求項25】 前記演算メカニズムが、前記走査デバイスによって前記表を読み取るように構成される請求項24に記載のデバイス。

【請求項26】 前記演算デバイスが、パームタイプコンピュータ及びワイヤレスネットワーク接続を含む請求項22に記載のデバイス。

【請求項27】 前記走査デバイスが、補助装置の遠隔制御に統合される請求項22に記載のデバイス。

【請求項28】 前記補助装置が、コンピュータ、テレビ及びウェブテレビデバイスのうちの一つである請求項27に記載のデバイス。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ビデオ又は他の記録媒体のある位置へのアクセス及びこの位置の検出に関する。本発明は、より具体的には、記録媒体内の位置を、その位置に関する参照マーカーを使用して参照することに関する。参照マーカーとしては、ノート、媒体自体の一部分又はその位置に関する他のアイテムを利用できる。また、本発明は、参照マーカーを記録媒体の対応する位置に関連付けるグラフィック符号化の使用にも関する。

【0002】

【従来の技術】マルチメディアノート作成システムは一般にミーティング中の音声及びビデオを取得し、取得した資料からスライドを形成する(例えば、ライブボード(LiveBoard)で実行されるシステムであるチボリ(Tivoli), Moran, T.P., Palen, L., Harrison, S., Chiu, P., Kimber, D., Minneman, S., vanMelle, W., and Zeilweger, P. による "I'll get that off the audio": a case of study of salvaging multimedia meeting records. Proceedings of CHI '97 CM, New York, pp.202-209を参照のこと)。チボリは、プレゼンテーションミーティングよりもワーキングミーティングをサポートするように設計されている。音声に索引付けされるチボリのインクストロークは、チボリスライド上の準備した資料と共に、ミーティングのグループノートになる。ラップトップを使用する参加者は、タイプしたテキストのコメントをチボリのスライド上に「発信」してよい。

【0003】類似する例のクラスルーム2000 (Classroom 2000)では、プレゼンテーションスライドの画像及び音声は取得されるが、ビデオは使用されない (Abowd, G.D., Atkeson, C.G., Brotherton, J., Enqvist, T., Gulley, P., and LeMon, J. Investigating the capture, integration and access problem of ubiquitous co

mputing in an educational setting, Proceedings of the CHI '98 Conference. ACM, New York, pp. 440-447; 及び Abowd, G.D., Atkeson, C.G., Feinstein, A., Hmelo, C., Kooper, R., Long, S., Sawhney, N., and Tani, M., Teaching and learning as multimedia authoring: the classroom 2000 project, Proceedings of the ACM Multimedia 796 Conference. ACM, New York, pp. 187-198を参照のこと)。さらに、クラスルーム2000では、発表者が標準図形形式のスライドを準備する努力が求められる。スライドはライブボードに表示され、スライドがプレロードされたPDA装置を用いてノート作成が行われる。これらのノートは、ライブボードの前で講義した教授によって注釈が付けられた音声及びスライドに後日同期される。

【0004】また別の例では、フォーラム (Forum) (Isaacs, E.A., Morris, T., and Rodriguez, T.K. A forum for supporting interactive presentations to distributed audiences. Proceedings of CSCW '94. ACM, New York, pp. 405-416を参照のこと)は、分散したプレゼンテーションの手段としてビデオを使用するシステムである。プレゼンテーション中、発言者を含む全員がワークステーションの前に座っている。特定の形式でスライドを準備しなければならない。テキスト及びマウスで描いたマークによってスライドに注釈を付けることもできるが、ビデオ画像には注釈を付けることができない。

【0005】別の例では、ストリームス (STREAMS) (Cruz, G., and Hill, R. Capturing and playing multimedia events with STREAMS. Proceedings of the ACM Multimedia '94 Conference. ACM, New York, pp. 193-200を参照のこと)は、室内カメラからのビデオを使用するプレゼンテーション取得のためのシステムである。これらのカメラはディスプレイ上の任意のプレゼンテーション内容を取得するためにも使用する。この方法は、室内の活動がディスプレイを不明瞭にするという問題がある。テキストのコメントを付加することによってレビュー中に取得したビデオストリームに注釈をつけることはできるが、プレゼンテーション中のノート作成はサポートされない。これらのどのシステムも、カメラからのライブ画像とプレゼンテーション資料をノートへ対話式に統合させることができない。

【0006】さらに、スタンドアローンインク及び音声ノート作成システムが幾つか知られている。例えば、FX PAL Dynamite (Wilcox, L.D., Schilit, B.N., and Sawhney, N. Dynamite: A Dynamically Organized Ink and Audio Notebook. Proceedings of CHI '97. ACM, New York, pp. 186-193を参照のこと); 及び音声録音に紙を使用するオーディオノートブック (Audio Notebook) (Stifelmann, L. The Audio Notebook: Paper and Pen Interaction with Structured Speech. Ph.D. Thesis. MIT, 1997を参照のこと)が挙げられる。ファイロチャット

(Filochat) (Whittaker, S., Hyland, P., and Wiley, M. Filochat: handwritten notes provide access to recorded conversations. Proceedings of CHI '94. ACM, New York, pp. 271-276を参照のこと)は、音声を手書きのノートで索引付けされる、ペンタブレットを伴うPCコンピュータである。また、ノータイム(NoTime) (Lanning, M., and Newman, W. Activity-based information technology in support of personal memory. Technical Report EPC-1991-103, Rank Xerox, EuroPARC, 1991)は、ユーザのインクストロークを録音した音声又はビデオに合わせるように設計された。

【0007】ビデオ注釈システムも知られている。マーキー(Marquee) (Weber, K., and Poon, A. Marquee: a tool for real-time video logging. Proceedings of CHI '94. ACM, New York, pp. 58-64を参照のこと)は、ビデオテープを見ながら注釈を付けるペンベースのシステムである。マーキーのより新しいバージョンは、WhereWereWeマルチメディアシステムからのデジタルビデオストリームにタイムスタンプを付けるという修正がなされている(Minneman, S., Harrison, S., Jassen, B., Kurtenbach, G., Moran, T., Smith, I., and van Melle, B. A confederation of tools for capturing and accessing collaborative activity. Proceedings of the ACM Multimedia '95 Conference. ACM, New York, pp. 523-534を参照のこと)。ヴァナ(Vanna) (Harrison, B., Baecker, R.M. Designing video annotation and analysis systems, Graphics Interface '92. Morgan-Kaufmann, pp. 157-166を参照のこと)及びEVA(MacKay, W. E. EVA: An experimental video annotator for symbolic analysis of video data, SIGCHI Bulletin, 21(2), 68-71. 1989. ACM Pressを参照のこと)は、テキストベースのシステムである。ビデオノーター(VideoNoter) (Trigg, R. Computer support for transcribing recorded activity, SIGCHI Bulletin, 21(2), 68-71. 1989. ACM Pressを参照のこと)は、アクティビティ(ビデオ、数字、ホワイトボードの図、テキスト)の異なるストリームを表示し、これらを同期させるが、音声からテキストを転記したりホワイトボードから図を抽出したりするために後生成を要する。これらのシステムは、デジタルビデオではなくビデオテープの使用に基づいた設計によって限定される。これらのどのシステムもビデオ画像のノートへの対話式統合ができない。シャープザウルス(Sharp Zaurus) (Zaurus Operation Manual. Sharp Corporation, 1996)は商品であり、デジタルカメラが付いたPDAである。デジタル写真を撮影し、手書きのノートにリンクさせることができる。

#### 【0008】

【発明が解決しようとする課題】さらに、多くのシステムがビデオ再生メカニズムを含む。図1は、ビデオディスプレイ領域102を含む従来のビデオ再生装置100

を例示する。タイムライン105は、表示されているビデオの現在の位置を示し、再生115、巻き戻し120、早送り125、コマ戻し130、コマ送り135等を含む標準VCRコントロールが設けられることでユーザがビデオをナビゲートできる。しかし、ビデオ内の特定の位置を検出する場合、特にノート又はビデオに関連する他のアイテムを参照してビデオ内の特定の位置を検出する際に問題が生じる。

#### 【0009】

【課題を解決するための手段】ノート作成中に監視されるビデオ又は他のデータストリームの重要部分が単純且つ正確な検索メカニズムで提供されるならば、ノート作成システムは再生しやすさが増し利用しやすくなることを、発明者は認識した。簡単にいうと、本発明は、ノート作成セッション(ミーティング、プレゼンテーション又は他の活動)の最中にユーザがキーフレームを選択したり注釈を付けることができるマルチメディアノート作成システムを提供する。各キーフレーム及び/又は注釈が取得され、ノート作成セッションのビデオ又はノート作成に関連する他のデータストリームの位置に索引付けされる。取得したキーフレーム及び注釈は符号化したグリフと共に紙にプリントされ、プリントされたシートは、ターゲットキーフレーム又は注釈に関連するグリフコードを走査し、ターゲットキーフレーム又は注釈によって索引付けされたビデオ又はデータストリームの重要な部分を参照することによってビデオの重要な部分にアクセスするために使用される。

【0010】これらの及び他の目的は、デジタルビデオ及びインクでノートを作成するシステム(ノートブックともいう)によって達成される。本発明は、以下に限定しないが以下のうちの何れかを含む多くの方法で具体化され得る。

【0011】データストリームの少なくとも一部分に対応する少なくとも一つのオブジェクト及びオブジェクトに対応するデータストリームの一部分を特定する各オブジェクトに関連するコードを含む出力を含む、データストリームにアクセスするためのインタフェースである。

【0012】データストリームに関連する少なくとも一つのオブジェクトを特定し、関連する各オブジェクトに対応するデータストリームのオフセットを決定し、各オフセット及び対応する関連オブジェクトを特定する索引を形成する、データストリームへのアクセス用の索引を形成する方法である。

【0013】データストリームに対応するオブジェクトを走査し、走査したオブジェクトに基づいてデータストリームのオフセットを決定し、決定したオフセットに対応するデータストリームの一部分にアクセスする、データストリームアクセス方法である。

【0014】ビデオ索引に埋め込まれたグリフコードを走査するように構成された走査デバイスと、走査したグ

リフコードを読み取り、グリフコードに基づいてビデオの一部分にアクセスするように構成された演算デバイスと、検索したビデオ部分を出力するように構成されたディスプレイメカニズムと、を含む、ノートからビデオにアクセスするデバイスである。

【0015】また、本発明は、ノート作成デバイスへの複数のデータストリームの入力、後に参照するために取得した部分又はフレームを有し、ユーザが取得した任意のデータストリーム（このようなデータストリームは、ペンベース入力、ビデオ入力、複数のビデオ入力、データストリーム、ホワイトボード入力などを含む任意のユーザ入力を含む）に関連する注釈を付けることができる方法も含む。この方法は、注釈を付けるか又はデータストリームの入力のうちの一つからフレームを選択し、注釈又は選択を格納して索引付けし、格納した選択又は注釈を特定するグリフを符号化し、付加することを含む。

【0016】プレゼンテーション又はデータ収集の最中に行った複数の選択及び／又は注釈に対してステップが繰り返される。注釈、選択及び付加したデータグリフは紙にプリントアウトされるか選択のための他の電子媒体に配されて保存した選択又は注釈の再生を始めるか又は注釈若しくは選択及び関連するグリフに関連するポイントでデータストリームを再生する。データグリフを走査するためにスキャナを利用してもよく、グリフから走査されたデータはビデオを特定し検索するために利用され、付加された（又は他の）デバイスは、走査したグリフに関連する選択及び／又は注釈によって特定されたデータストリームの位置からビデオ又は他のデータストリームの再生を開始する。

【0017】本発明を自動的に実施することができ、これによってデータストリームが取得され、キーフレーム又は他のデータ選択を特定することによって自動的に要約され、このキーフレーム又は他のデータ選択は抽出され、選択が抽出されたデータストリームの位置を特定するグリフが埋め込まれる。ユーザによって手作業で付けられた注釈及び／又は行われた選択を参照するグリフは、別個に形成されてもよいし、選択したキーフレームの出力に含まれていてもよい。

【0018】本発明は、ネットワーク、例えばデバイスが注釈又はノート作成のためにダウンロード又はブロードキャストビデオストリームを取得するインターネットで実施され得る。本発明の出力、例えば選択及び関連するグリフを有する紙出力はユーザによって走査され、グリフはhttp又はサーバーへの他のネットワーク関連アドレスを特定し、サーバーは走査したグリフによって示された位置からビデオを再生し、PC、パームデバイス、ウェブテレビベース又は他のディスプレイメカニズムに出力する。

【0019】注釈又はプリントページの画像に関連する

ブロックグリフを含む本発明の出力の種々の実施の形態が説明され、グリフのアドレスカーベットが注釈付きタイムラインに関連して提供され、時間にマッピングされたX-Y位置を有する二次元アドレスカーベットは、任意のデバイスの検索インタフェース又はウェブブラウザタイプのプロダクトのサマリーインタフェースに類似するビデオアクセスへの選択したアクセスを提供する。

【0020】

【発明の実施の形態】以下の詳細な説明を参照し、添付図面に関連して考慮することによって本発明をより理解するにつれて、本発明のより完全な価値及び本発明の利点の多くを容易に得ることができであろう。

【0021】同一参照番号が同一又は対応する部分を示す図面を再び参照すると、具体的には図2を参照すると、ノート作成アプリケーション210を実行するペンベースのコンピュータ200が例示されている。ノート作成アプリケーションは、画像を書き込み、取得し、注釈を付けるための紙のノートのページに似たメイン領域215を含む。アクティブビデオを見るための小さなビデオウィンドウ220がある。アクティブビデオはアクティブチャンネルから得られ、ユーザはチャンネルを変更して他のビデオストリームを見たり、別のタイプのデータストリームをユーザ又はアプリケーションに表示することができる。ノート作成アプリケーションは、ビデオウィンドウに示されているフレームを取得し、取得したフレーム（例えば、225）をメイン領域のマージンに配置したり、フレームを大きな背景画像としてセットで取得するための設備を含む。ユーザのペンデバイス230の「インク」ストロークも取得される。サムネイル、背景画像及びインクストロークにはタイムスタンプが付けられ、ビデオ又は他のデータストリームの索引を提供するために利用される。

【0022】アクティブビデオは、あらゆるソースを介してノート作成アプリケーションに出力されることができる。ビデオソースは、種々の方法で、例えば室内カメラ又はドキュメントカメラから、リアプロジェクター、テレビ、オーバーヘッドスライドのドキュメントカメラ、VCR又は任意のビデオストリームへのタップから、又はペンコンピュータに装着された小さなポータブルカメラから、又は任意の他のタイプのデータ入力デバイスから取得することができる。多数のビデオ又は他のデータソースがノート作成システムに出力される。

【0023】図3及び4は、ビデオウィンドウ320及びビデオコントロール310を含むノート作成システムのスクリーンショット300を例示する。ビデオウィンドウの隣にはビデオと対話するための三つのボタンがある。即ち、一番上のボタン（チャンネルチェンジャーボタン322）はチャンネルを変更し、中間のボタン324はサムネイル（例えば、サムネイル370）をノートページ領域315のマージンに撮影し、一番下のボタン

326は、大きな画像390（即ち、注釈に利用するか或いは見るために拡大したかもしれない画像）をノートページ領域380に撮影する。

【0024】ビデオウィンドウ320のさらに左側はVCR型コントロール360のセットである。タイムライン350は、再生されているビデオの現在の時間位置を示すポインタ355を有する。また、ポインタは、ノート作成セッションの最中に取得されている多数のデータストリーム（例えば、ビデオ入力、ホワイトボード入力及びユーザからのペンベース入力の各々）の時間位置も示す。また、タイムラインは、ブラウジング及びレビューを容易にするために索引付け情報も表示する。ノート作成デバイスの正確な構造は任意の形態であってよく、本明細書中に説明したものと異なるメカニズム又は別の制御メカニズムを用いてもよい。

【0025】このシステムによって、ユーザはデジタル化されサーバに格納されている（又は局所的に格納されているか又はテープ、CD等のような他のデバイスに格納されている）ビデオを見ながらペンベースのコンピュータでノートを取ることができる。ノートにはタイムスタンプが付けられ、ビデオに同期される。さらに、ユーザは、ビデオからの静止画像を、マージンのサムネイルとして又はノートページのフルサイズの画像として、ノートに撮影することができる。これらの画像にもタイムスタンプが付けられ、またビデオに同期されているため、索引の役目も果たす。

【0026】ノート作成セッションの後、録画したビデオを再生するためにノート作成システムを使用することができる。標準VCRコントロール、例えば早送り又は巻き戻しによって再生位置を制御することができる。さらに、ペンストローク及びノートの静止画像によって形成された索引を使用してビデオの特定の部分にアクセスすることができる。ユーザは所望の索引オブジェクトを選択し、再生ボタンを押す（又は例えばダブルクリック（Double Click）等のような他の再生メカニズムを用いる）。ビデオ再生時間を示すために注釈付きタイムラインも利用できる（図1を参照）。

【0027】上記のノート作成システムは、FX Palo Alto Laboratoryで開発したノートルック（NoteLook）システムについて簡単に説明している。本発明を適用する画像、スナップ、注釈などを取得するためにノートルック及び他のノート作成システムを利用することができる。本発明を適用するために、ビデオ、音声、マルチメディア及び他のデータタイプに対する取得、注釈、強調又は他の処理をサポートする多くの他のタイプのデバイスを利用することができる。必要なのは、任意のタイプのデータストリームを取得できるデバイス及び本発明の方法及び／又は処理を適用するための（デバイスに含まれるか又はアクセス可能な）設備のみである。

【0028】帯域幅を節約するためにフレームレート圧

縮を利用するのが好ましい。自動変更検出（例えば、スライド変更検出）を利用してもよく、メディアストリームにおける変更のみがノート作成デバイスに送られる。或いは、メディアストリームにおけるマテリアルの秒当たりのフレームレート圧縮を減少させることもできる（即ち、30fpsのビデオショットを1又は2fpsに減少させてもよい）。これによってノートルックのクライアントでのディスプレイにジッターが現れるかもしれないが、ノート作成デバイスでの完全なビデオはノート作成処理においては必須ではなく、実際、ノート作成者がビデオに出席している状況では（ノート作成者は何が行われているかを見聞きできるため）必要ない。ノート作成デバイスは、ビデオ／データストリームを格納してもよいし、ビデオ／データストリームは別のサーバーに取得又は維持されてもよい。

【0029】ノート作成システムは、室内カメラ及びプロジェクターからのビデオをワイヤード又はワイヤレスネットワークを介してノートルックコンピュータに伝送できる会議室の設定で特に有用である。このようなノート作成デバイスを会議室で利用することは簡単に想像できるが、ユーザが室外で使用するためのデバイスを個々に所有することはあまり想像できない。したがって、会議室外でノート及び録画したビデオにアクセスするための別の手段をユーザに与えることが重要である。このための最も簡単な方法の一つは、ユーザにノートのプリントアウトを配布することである。しかし、これだけでは録画したビデオへのアクセスは与えられない。

【0030】本発明は、ノート、注釈、クリップ、スナップショット（集散的にノートと称する）のデータグリフを、ノート作成セッションの最中に利用したビデオデータストリームへのアクセスを提供する他のプリントアウト又は他の出力に埋め込む。グリフはビデオの時間のオフセットを提供するため、ビデオの特定の部分にアクセスすることができる。さらに、グリフはこれらのノートに関連する録画したビデオの位置を符号化することができる。

【0031】紙からビデオへのアクセスを提供するためにグリフを使用できる別のアプリケーションはビデオサマリーである。プリントに適したビデオサマリーの一つのタイプは、1999年3月12日に出願した出願番号第09/276,529号の“AUTOMATIC VIDEOSUMMARIZATION USING A MEASURE OF SHOT IMPORTANCE AND A FRAME-PACKING METHOD”というタイトルのUchihashiによる（全て参照として本明細書中に援用される）漫画形式のサマリーである。ここで、ビデオは自動的に分析され、サイズ変更可能なキーフレームがサマリーとして選択される。図4は、ビデオサマリーの例を示す。大きなキーフレームはビデオのより重要な部分に対応する。ビデオへのアクセスを提供するウェブページにサマリーを配置することができる。キーフレームが選択されると、ビデオはそのフレームか

ら再生を開始する。

【0032】ある実施の形態では、サマリーは埋め込まれたグリフと共にプリントされる。これによって、ウェブ上のコンピュータインタフェースを介してではなく、紙からビデオへのアクセスが可能となる。ユーザはハンドスキャナを使用してプリントしたビデオサマリーの選択した部分に埋め込まれたグリフを走査することができる。スキャナはウェブテレビ(WebTV)のようなデバイス又は通常のPCに接続され、これらは走査したグリフによって索引付けされたビデオ/データストリーム部分にアクセスする。

【0033】プリントされたノート又はビデオサマリーに埋め込まれたデータグリフによって、ユーザはビデオにインタフェースできる。ビデオは、ビデオの位置、ビデオのセグメント、少なくとも一つのマルチメディアのセット又は他のデータストリームを参照する埋め込まれたグリフを有するビデオから取得したノート又は画像を選択することによってアクセスされる。

【0034】図5は本発明のデバイスを例示する。取得デバイス400は、音声、ビデオ、ホワイトボード又は他のデータ入力の何れでもよい少なくとも一つの入力データストリーム410を含む。ユーザ入力420は、任意のデータ入力ストリーム410の部分を選定し、入力ストリーム又はユーザがノートを取った他のイベントに関連する注釈を付ける及び/又はノートを取るために利用される。

【0035】取得デバイス400は取得/注釈メカニズム425を含み、このメカニズムは、入力データストリーム410のユーザ選択部分及びユーザが付けた注釈又はユーザがとったノートを取得し、これらのユーザ選択部分、注釈及び/又はノートにID、索引又は入力データストリームにおける位置を選定するタイムスタンプを付け、これらを格納デバイス430に格納する。取得デバイス400は、入力データストリームを格納のために格納デバイス430に送るためのチャンネルも含む。

(或いは、入力データストリームを独立的に格納デバイス430に出力してもよいし、単にサーバーで保持してもよい。)

【0036】サマリーメカニズム440は取得したスナップショット、注釈、ノートを検索し、これらを出力に適した形式にフォーマットする。さらに、サマリーデバイス440は、独立した要約化及び一つ以上の入力データストリーム410のキーフレーム選択を提供し得る。出力は、パーソナルコンピュータ450のブラウザデバイスのようなウェブディスプレイ用にフォーマットされるか、ハードコピー形式(紙)455にプリントアウトされるか、電子形式(例えば、フロッピー(登録商標)ディスク)460に格納される。出力は、他のウェブベースの製品、例えばウェブテレビ等を含む他のタイプの出力デバイスに出力されてもよい。

【0037】図6は本発明の処理を例示する。ステップ500では、ノート作成システムがスタートする。ノート作成システムは、上記に示した特徴のうちの一つ以上を含むか又はユーザによってコメントされ、注釈が付けられ、選択されたデータストリームを受け取る任意のタイプのデバイスである。ステップ510でノート作成システムのユーザは、注釈を付けるか又はノート作成システムに出力されるデータ部分を選択する。ステップ520で、付けた注釈又は行った選択を保存し、選択及び/又は注釈の検索ができるように索引付けし、ノート作成システムに出力されるデータストリームの少なくとも一つの位置を選定する。ノート作成システムには出力されないがノート作成システムに関連して再生されるデータストリームを参照して注釈を付けることもできる。

【0038】ステップ530で、注釈又は選択のいずれか及び特定の一個或いは複数のデータストリームの位置を選定する情報と共にグリフが埋め込まれ、グリフが選択又は注釈に付加される(又は参照として維持される)。ステップ540で、注釈/選択処理及びグリフ付加が繰り返される。

【0039】当業者に理解されるように、本発明の開示に基づいて、図6のフローチャートの多くの異なるバリエーションを本発明を実施するために実行してもよい。例えば、選択及び注釈の各々をバッチ形式で符号化する前に、ループで選択及び注釈を全て行い、格納してもよい。処理は、アーキテクチャ及びプログラミング技術に応じてパラレル又はシリアルに実行することができる。重要なのは、注釈が付けられ選択が行われ、グリフが注釈及び/又は選択の各々に関連付けられることであり、グリフは、注釈及び/又は選択に対応するデータストリームの一部分を選定する情報を含むか該情報を指示する。

【0040】グリフは、(特定のノートセットに対応する)サーバー上の関連するビデオの位置の情報又はそのビデオの時間のオフセットを含んでもよいし、他のタイプの写真、ビデオ、音声、文書又はデータのポイントを含んでもよい。選択が行われるとき又は注釈が付けられるときに表示されているビデオにリンクするために、又はそのとき注釈が付けられているかコメントされている文書にグリフをリンクさせるためにグリフを使用するのが最も理にかなっている。或いは、ユーザは(ポップアップウィンドウ又はノート作成デバイスによって提供される他のメカニズムによって)ビデオの位置を手作業で選定する。

【0041】走査デバイス(例えば、ハンドヘルドスキャナ)を使用して紙のノートに出力すると、ユーザは紙のノートの画像領域を走査し、これによって所望の選択を指定する。走査画像のグリフによって適切な情報が与えられ、特定の時間のオフセットでビデオの再生を開始する。

【0042】グリフの種々の実施の形態が本発明で利用される。一つの実施の形態では、プリントアウト又は他の出力の各索引付きオブジェクトに対してブロックグリフを使用する。図7の例には、ノート作成デバイスのページ上に、サムネイル（スナップ）610に関連するグリフ615及びインクストローク（インクストローク620、630及びグリフ625、635）の各グループが示されている。オブジェクト（サムネイル及びインクストローク）を、図6の処理又はこれと同等の処理によって索引付けし、各グリフに付加する。他のタイプのオブジェクトも同様に索引付けされてもよく、背景画像を含む画像又はデータを含むグリフが付加されてもよい。

【0043】ある実施の形態では、オブジェクトに関連するビデオ（又は他のマルチメディア／データストリーム）の位置は、ノートページ上の付加されたグリフで符号化される。或いは、グリフは、以下に説明するようにビデオ及びオフセットを特定する表へのエントリーポイントを提供してもよい。

【0044】インクストローク自体は関連するインクストロークのグループを表す。各ペンストロークはビデオに索引付けられるのではなく、本発明はペンストロークのグループを索引付けする。ペンストロークをグループ化する多くの方法を利用することができ、これらの方法は本発明の範囲内である。例えば、構文構造によるグループ化（例えば文字認識及び文法の規則を使用して完全な文又はパラグラフをグループ化する）、単語認識によるグループ化（単語をグループ化する）又は時間によるグループ化が挙げられる。しかし、ペンストロークは時間によってグループ化されるのが好ましいため、グループのストローク間の時間差が小さい。したがって、索引付け解像度のいかなるロスの重要性も減少される。1秒間隔がペンストロークをグループ化するのに最もよいことが本発明者によって判定された。しかしながら、本発明はこれに限定されない。

【0045】このデザインは、ビデオサマリー又は選択、注釈及びビデオサマリーの任意の組み合わせにも通用する。ビデオサマリーの場合、ブロックグリフはサマリーの各キーフレームに関連する。

【0046】図8は、ビデオを要約しグリフを埋め込む本発明の一つの実施の形態の処理のフローチャートである。ステップ700でビデオを要約する。ビデオは、ノート作成デバイスに入力されたビデオストリームであってもよいし、会議、ミーティング、活動又は他のイベントに関連して個々に録画した任意のビデオストリームであってもよい。ビデオは、例えば音声データストリームのように、要約処理が施されたオブジェクトである必要はなく、他のタイプのプロダクトを要約し索引付けしてもよい。視覚画像で通常表されるプロダクト以外のプロダクト（例えば、音声）の場合は、アイコン、グラフ又はプロダクトを特定する他の表示が含まれていてもよい。

い。

【0047】ステップ710で要約したビデオからキーフレームを選択する。任意の選択アルゴリズムを利用することができ、及び／又は外部基準のセットを適用してキーフレームを選択することもできる。ステップ720でキーフレームをビデオ又は他のデータストリームに索引付けし、索引付き位置を符号化するグリフに関連付ける。他のデータは、ビデオ又はデータストリーム、http、ftp又は下位のデータストリームがアクセスできる他のアドレスの特定を含むグリフに符号化することもできる。或いは、下位のデータストリームがアクセスできる表エントリーを参照するロケータ番号又は他の識別子でグリフを符号化してもよい（即ち、ステップ720は参照表の構築を含んでもよい）。

【0048】或いは、自動的に取得され（例えば、変更検出基準を使用して）特定のノートに形成された重要なビデオショット（例えば、ホワイトボードデータストリームのスライド変更）を自動的に決定するために、自動索引付けを利用することもできる。このような自動索引付き資料を、図9の例えばサマリーに示されるスライド変更810及び25:00の対応するタイムラインスナップショット860のように、タイムラインにリンクさせてもよい。

【0049】ステップ730では、他の注釈又は選択も索引付けされ、対応するグリフが提供される。ステップ740では、出力（例えば、紙）が生成され、この出力からグリフによって索引付けされたビデオ及び／又はデータストリームを参照したり呼び出したりする。

【0050】他の実施の形態では、埋め込まれたデータグリフがグリフアドレスカーベットの形式で提供される。例えば、グリフアドレスカーベットは、プリントされたノートページ又はサマリーの下の方の水平なタイムラインに配置される。図9は、グリフアドレスカーベット820及び関連するタイムライン840を有するビデオサマリー800を例示する。タイムライン840は水平位置を時間にマッピングするために使用され、グリフアドレスカーベット820の対応する位置に線形又は非線形に符号化されてビデオにおいて索引付けされているため、ユーザはその時間に（ノートからの）索引オブジェクト又は（サマリーからの）キーフレームがあるか否かに関わらず、ビデオの任意のポイントにアクセスすることができる。

【0051】強調として、ナビゲーションを助けるために、（ビデオサマリー又はノートからの）任意のオブジェクトをタイムライン上の適切な位置に表示してもよい。図9には種々のサマリーオブジェクト（キーフレーム）860が示され、ビデオのサマリーオブジェクトの位置に対応するグリフアドレスカーベット820の符号化された索引に対応する位置に配置されている。

【0052】データグリフカーベットの配置は、ユーザ

がキーフレームオブジェクトとタイムラインとの関連を知覚できるようにタイムラインに関連付けさえすればよい。タイムライン上のオブジェクト又は位置が走査されるとそのタイムライン上のオブジェクト又は位置に対応するグリフコードも走査される。したがって、グリフアドレスカーペット820の位置は図9に示すようにタイムラインより上であってもよいし、タイムラインの下でもよいし、タイムラインと統合されていてもよい。

【0053】線形又は非線形タイムスケールを表示するために、タイムライン又はグリフアドレスカーペットに注釈及び／又は着色が施されてもよい。同一の又は離散的な(個別の又は異なる)タイムスケールで多数のビデオ(又は他のデータ)チャンネルにアクセスするために、多数のパラレルアドレスカーペットストリップを使用することができる。多数のビデオチャンネルが記録され撮影されるノートブックのようなノート作成デバイスを利用する場合、これは有用である。ビデオの長いタイムスケールに分散した選択を与えるために、多数のストリップを使用することもできる。

【0054】分散した選択によって、ビデオ又は他のデータストリームの部分が選択され、移動され、他のファイルに付加される。例えば、グリフデータカーペットと共に実行するビデオタイムラインでは、タイムライン(及び埋め込まれたグリフ)の部分が選択されてもよい。これらの選択はドラッグされ、他のアプリケーションにドロップされ、別個のファイルとして格納され、再生されることができ、任意の機能が実行される。これらの機能は、その機能を実行するために構成されたアプリケーションを表すアイコンに選択をドロップするか又は例えばポップアップウィンドウ(例えば右クリックポップアップウィンドウ)からさらに選択することによって開始する。機能又はアプリケーションは、選択によって表されるビデオクリップで動作する。

【0055】タイムラインのアドレスカーペット(又は本明細書中に説明した他の実施の形態)の場合、アドレスカーペット上の水平位置をビデオに関連する時間にマ

ッピングするために、表又は他の機能的マッピングを利用できる。この表は紙の上のグリフコードとして格納され、このグリフコードは走査され演算メカニズムに読み取られる。該演算メカニズムは、連続的に走査されたグリフカーペットの任意の部分を読み取られた表に対して参照して走査したグリフカーペットの部分に対応するビデオの索引を決定する。或いは、表をサーバー(例えば、デジタル化されたビデオ又は他のデータも維持するサーバ)に格納してもよい。ビデオの位置は、別のブロックグリフによって提供されることができ。

【0056】他の解決策は、ビデオ位置の分散したラベルをタイムラインのアドレスカーペットに統合することである。例えば、ビデオ名(ID)及びビデオの時間(位置又は索引)をタイムライン又は他のグリフに符号化してもよい。

【0057】別の実施の形態では、グリフ情報は二次元アドレスカーペットに埋め込まれる。図10は、二次元アドレスカーペット900の一つの実施の形態を例示する。グリフアドレスカーペット位置は1グリフの精度で選択され、この精度は一般的にページ上で約1/16インチであり、現在の電子コンピュータディスプレイの解像度と同等である。新しい、又は改良したグリフタイプの技術を使用する他の解像度を利用してもよい。本発明の最初の実施の形態はノート及び埋め込みグリフとの紙インタフェース(プリントアウト)を使用するが、他の実施の形態はあらゆるタイプの出力を利用してよい。例えば、ユーザによって走査されるノート及びグリフ符号化を表示するのに十分な解像度を有するコンピュータディスプレイがある。

【0058】幾つかの実施の形態では、表が形成され、ノートオブジェクト(ストローク、スナップ)の近隣の二次元カーペットのX-Y位置をビデオの時間にマッピングする。表1は、ノートオブジェクトのX-Y位置のマッピングの例示的な表を示す。

【表1】

X-Y位置	データストリームID	索引(タイミング位置)
(1, 4.5)	ビデオ1	2252
(6.5, 4)	インク	1.1
(4.25, 4)	インク	1.2

各X-Y位置はスナップ、インクストローク又は他のオブジェクトに対応する。ノート又は他のオブジェクトに対応するビデオにアクセスする選択メカニズムは、走査された二次元グリフの位置に最も近いノート又はオブジェクトを決定する。

【0059】或いは、表はより分かりやすく、オブジェクトに関連するビデオ位置にアクセスするための完全な範囲(二次元カーペット上の領域)を画定するかもしれない。例えば、スナップ910に関連するビデオは、ボ

ックス915によって示されるように(0, 0)～(2, 0)と(0, 5.5)～(2, 5.5)の範囲からアクセスされ得る。インクストローク920に関連するビデオは、925で示される境界内のグリフを走査することによってアクセスされ、またインクストローク930に関連するビデオは935で示される境界のグリフを走査することによってアクセスされ得る。境界のサイズ及び形状は、アクセスされるオブジェクトに関する領域を分割する任意の方法によって演算されることができ

る。

【0060】タイムラインの場合と同様に、グリフコードをビデオ又は他のデータストリームの索引に関連させた表は、紙の上のグリフコードとして格納されるか又はデジタル化されたビデオと共にサーバーに格納されるような他の構造として格納されることができる。或いは、ビデオにアクセスするために必要な全ての情報と共にグリフ自体が符号化される。

【0061】他の表では、ビデオストリーム及び索引（又はビデオのオフセット）をX-Y位置によって参照する代わりに、グリフに埋め込まれたコードが走査デバイスによって読み取られ、ビデオ及びビデオのオフセット又は索引を参照するために使用される。

【0062】ビデオの位置を、拡張アドレス分散ラベルとして、又はインターリーブもしくはインターレースデータコードとして、アドレスコードに埋め込むことができる。例えば、グリフに埋め込まれたデータ、例えばビデオID及び位置は、埋め込まれたグリフコードの他の部分で符号化されてもよい。或いは、グリフコードのデータを暗号化又は符号化する他の方法を利用してもよく、この方法は、ヘヒト(Hecht)による"Parallel Propagating Embedded Binary Sequences For Parameterizing Two Dimensional Image Domain Code Patterns in Two Dimensional Space"というタイトルの米国特許第5,825,933号（全体として参照によって本明細書中に援用する）に開示したような技術を含む。二次元アドレスカーベットのアプリケーションはビデオサマリーのアプリケーションに同様に適用される。

【0063】図11は、本発明をネットワーク環境で適用するためのハードウェアの一つの実施の形態を例示する。演算デバイス1000は、ノート作成セッション中に取得又は再生されるビデオ及び/又はデータストリームを維持する格納デバイス1010を含む。サーバープログラム1020は、ブラウザ又は他のタイプのデバイスからの要求に応じて取得した任意のデータストリームの部分を出力するように構成されている。演算デバイス1000及びサーバープログラム1020はそれぞれネットワーク1030（例えば、インターネット）に接続されている。ユーザは、周辺走査メカニズム1060及びディスプレイデバイス1055を伴う演算メカニズム1050を有する。演算メカニズム1050は、（1040を介して）ネットワーク1030に接続する標準パーソナルコンピュータ又はウェブテレビベースのデバイスである。一般的な実施の形態では、ネットワーク1030への接続1040は通信会社1035によって実行され、このような接続はワイヤード又はワイヤレス接続の何れでもよい。

【0064】本発明の1070からの出力は、種々のスナップ又はノート及び対応するグリフを含む。走査デバイスメカニズム1060を利用するユーザは、特定のノ

ート/スナップに関連するグリフを選択し、そのグリフを走査する。演算メカニズム1050は走査したグリフコードを読み取り、適切なメッセージをフォーマットし、このメッセージはネットワーク1030に送られサーバープログラム1020で受け取られる。サーバープログラムは、走査したグリフに対応する、格納デバイス1010に格納されたデータストリームの関連部分にアクセスする。アクセス部分はネットワーク1030を介して演算メカニズム1050に伝送され、ディスプレイデバイス1055（コンピュートースクリーン又は標準テレビセット）に表示される。

【0065】演算メカニズム1000、サーバープログラム1020及び格納デバイス1010は、演算メカニズム1000と同じ演算メカニズムに配置されていてもよいし、ネットワーク1030又は他のデバイスを介して接続される異なる演算メカニズムに保持されていてもよい。サーバー、演算メカニズム及びネットワーク1030の接続は、ワイヤード又は衛星、マイクロウェーブ、セルラー又は他のシステムを含むワイヤレス通信システムの何れでもよい。

【0066】演算メカニズム1050及びディスプレイデバイス1055は、モバイル演算デバイス及びディスプレイメカニズム、例えばラップトップコンピュータ、パームタイプコンピュータ又はグリフコードを入力するためのスキャナ若しくは他のメカニズムを含む任意の電子デバイスの形態を取ってもよい。演算デバイス1050及びディスプレイ1055は、車両例えば自動車、配達もしくは修理バンのダッシュボード又はホテルのロビー、会議室に隣接するホール又はシステムを利用できるあらゆる場所に備えられてよい。

【0067】ある実施の形態では、走査デバイスメカニズム1060がテレビ及び/又はウェブテレビデバイスの遠隔制御に統合されている。したがって、ユーザは、プリントされたサマリー又は記事、広告又は本発明によって埋め込まれたグリフを有する任意のタイプの媒体を有することができ、スキャナを利用してプリントされたサマリー、記事等に関連するビデオ又は他のデータにアクセスする。

【0068】本発明は、グリフ技術を利用してプリントしたノート及びサマリーからビデオにアクセスする。さらに、本発明は、表及び直接的な符号化を含む、ビデオ位置及び時間のオフセットをグリフに符号化する特定の技術を含む。

【0069】本発明は、グリフを使用して電子ノート作成デバイス（例えば、ノートブック）から取得したノートのプリントアウトからビデオへのアクセスを提供し、以下を含む。

- ・ブロックグリフをノートページの手書きストローク及び画像に関連付ける。
- ・グリフアドレスカーベットを使用してビデオのタイム

ラインを形成し、これによって紙のタイムラインからのビデオアクセスを提供する。

・ノートページにわたって二次元アドレスカーベットを使用し、これによって索引オブジェクトからビデオにアクセスするペンベースの走査デバイスのようなスキャナによって動作する紙インタフェースを提供する。紙インタフェースは図9のような出力をグリフカーベットページにプリントすることによって生成される。

【0070】プリントした漫画ビデオサマリーの場合は、図4に例示するように、

・グリフアドレスカーベットを使用してビデオのタイムラインを形成し、これによってウェブのタイムラインからのビデオアクセスに類似する紙からのビデオアクセスを提供する。

・サマリーページにわたって二次元アドレスカーベットを使用し、これによってキーフレームからビデオにアクセスするウェブインタフェースに類似する紙インタフェースを提供する。

【0071】また、本発明はサーバー上のビデオの位置及びビデオの時間のオフセットを特定する技術も提供する。これらによってビデオソース及び時間のオフセットが同じグリフに符号化されるため、ビデオソースを特定するのに別個の走査が必要ない。本発明は、以下を含む。

・ビデオのソース及び時間位置を特定するためのグリフカーベットアドレスにおける分散ラベルの使用。

・ソース及び時間位置を特定するためのアドレスカーベットにおけるインターレースコードの使用

【0072】実験結果によって、ユーザはタイムラインとオブジェクトベースのアクセスとの組み合わせを好むことが分かった。さらに、ビデオにアクセスするために紙を使用することが受容され、ビデオサマリーにも同様に有用であることが分かった。手書きのノートセグメント及びビデオフレームプリントからのアドレスをロバスト復号することで、ノートブックのノートページのオーバープリントがグリフカーベットページで行われてきた。

【0073】コンピュータ技術者には明らかであるように、従来の汎用又は専用デジタルコンピュータ又は本開示の教示によってプログラミングしたマイクロプロセッサを使用して、本発明を容易に実行してよい。

【0074】ソフトウェア技術者に明らかであるように、本開示の教示に基づいて、適切なソフトウェアコーディングを熟練のプログラマーによって簡単に準備することができる。当業者に明らかであるように、アプリケーション特定集積回路を準備することによって、又は従来の構成回路の適切なネットワークの相互接続によって、本発明を実施してもよい。

【0075】本発明は、本発明の処理の何れかを実行するためにコンピュータをプログラミングするのに使用で

きるインストラクションを格納した単数又は複数の記憶媒体であるコンピュータプログラム製品を含む。記憶媒体としては、以下に限定しないが、フロッピーディスク、光ディスク、DVD、CD-ROM、マイクロドライブ及び光磁気ディスクを含む任意のタイプのディスク、ROM、RAM、EPROM、EEPROM、DRAM、VRAM、フラッシュメモリデバイス、磁気又は光カード、(分子メモリICを含む)ナノシステム又はインストラクション及び/又はデータの格納に適した任意のタイプの媒体若しくはデバイスを含む。

【0076】単数又は複数のコンピュータ読取可能媒体のうちの一つの格納されて、本発明は、汎用/専用コンピュータのハードウェア又はマイクロプロセッサを制御するソフトウェア及びコンピュータ又はマイクロプロセッサが人間のユーザ又は本発明の結果を利用する他のメカニズムと対話できるソフトウェアを含む。このようなソフトウェアとしては、以下に限定しないが、デバイスドライバ、オペレーティングシステム、及びユーザアプリケーションを含む。最終的に、このようなコンピュータ読取可能媒体は上記のように本発明を実行するソフトウェアを含む。

【0077】汎用/専用コンピュータ又はマイクロプロセッサのプログラミング(ソフトウェア)に含まれるのは、本発明の教示を実施するソフトウェアモジュールであり、以下に限定しないが、メディアストリームの取得及び注釈付け、重要なノート作成イベントのタイムラインの生成、静止フレームのメディアストリームのポイント又はセグメントへのリンク及びこれを表示するグリフコードへのリンク、スライド変更の認識、キーフレームの要約、フレーム、インクストローク、及び他のデータをグリフ位置に相関させる表の作成、本発明の処理による結果の通信である。

【0078】本発明の多数の変更及びバリエーションが上記の教示によって可能であることは明らかである。したがって、本発明の請求の範囲内で、本明細書に具体的に述べた以外の方法で本発明を実施できることが理解されるであろう。

#### 【図面の簡単な説明】

【図1】従来のビデオ再生デバイスのスクリーンショットを示す図である。

【図2】ペンベースのコンピュータで実行するノート作成アプリケーションを示す図である。

【図3】ビデオ、ホワイトボード入力及びユーザ注釈領域を有するノート作成アプリケーションのスクリーンショットを示す図である。

【図4】本発明の漫画ビデオサマリーの例を示す図である。

【図5】本発明のビデオサマリーを生成するための注釈/データセグメント取得デバイス及び情報の流れの一つの実施の形態のブロック図である。

【図6】本発明の選択及び注釈の取得及び索引付けの処理を説明するフローチャートである。

【図7】本発明のグリフ付き注釈及び索引付きビデオセグメントの例示的な出力を示す図である。

【図8】ビデオを要約しグリフを埋め込む本発明の処理を説明するフローチャートである。

【図9】マルチメディアデータストリームのポイントにアクセスするためにタイムラインが付加され、関連するデータグリフカーベットを有する（ビデオ及びホワイトボードスナップを含む）キーフレームサマリーを例示する図である。

【図10】選択したキーフレーム及びノート作成システム

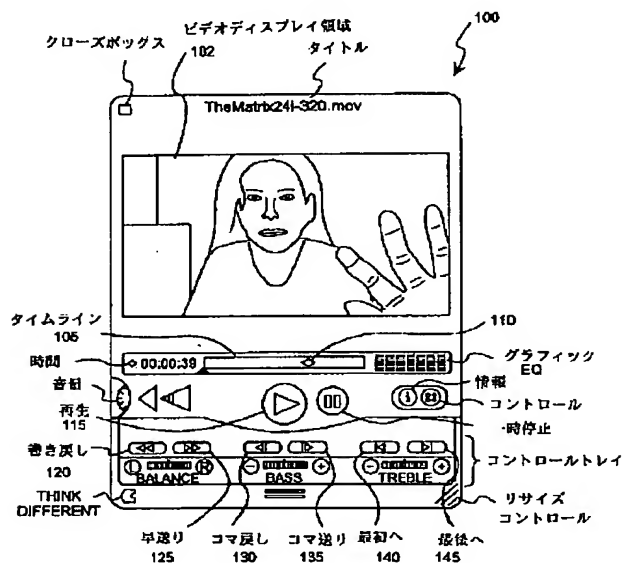
ムで取得した注釈の多数のポイントをマルチメディアデータストリームにマッピングする例示的なデータカーベットを例示する図である。

【図11】本発明を使用してネットワーク接続デバイスでビデオストリームのポイントにアクセスするために利用されるハードウェア構造のブロック図である。

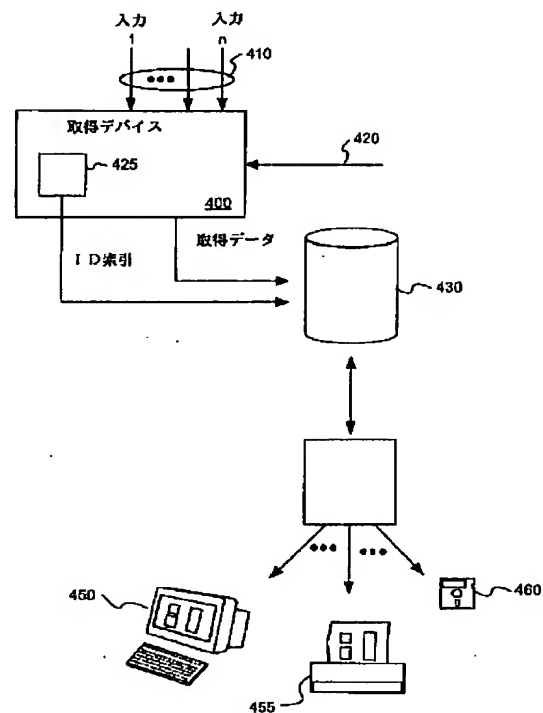
【符号の説明】

350、840 タイムライン  
615、625、635 グリフ  
800 ビデオサマリー  
820 グリフアドレスカーベット  
900 二次元アドレスカーベット

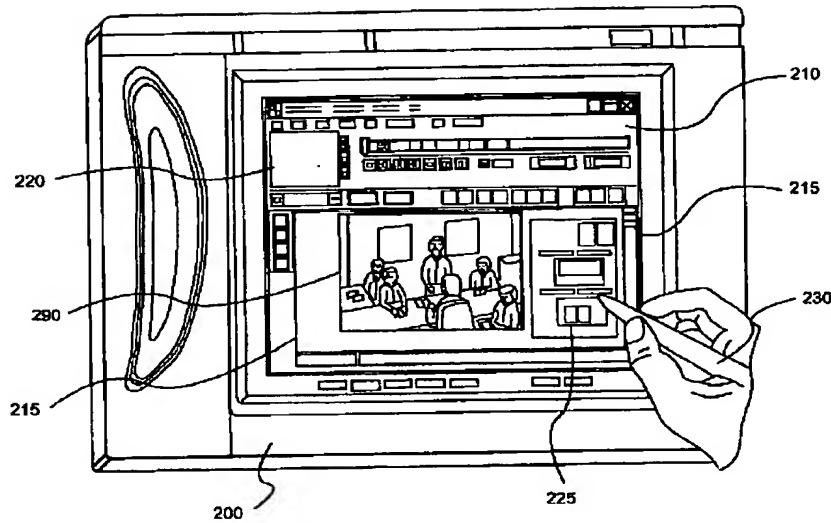
【図1】



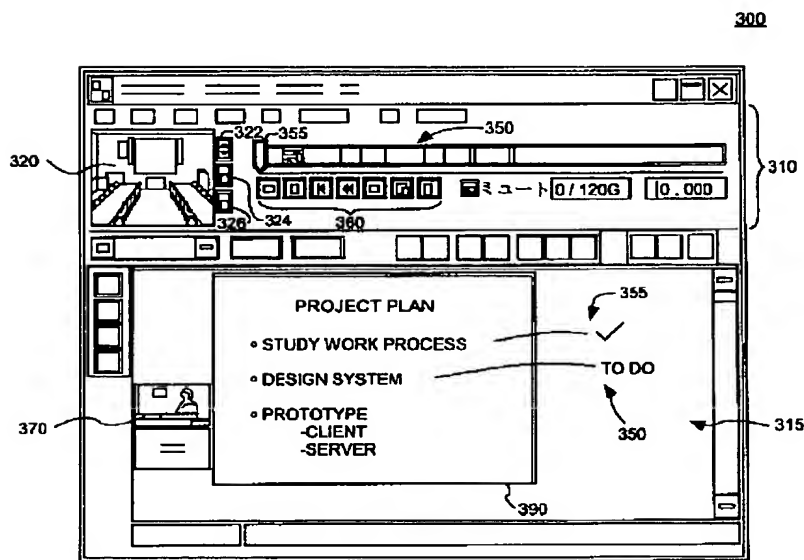
【図5】



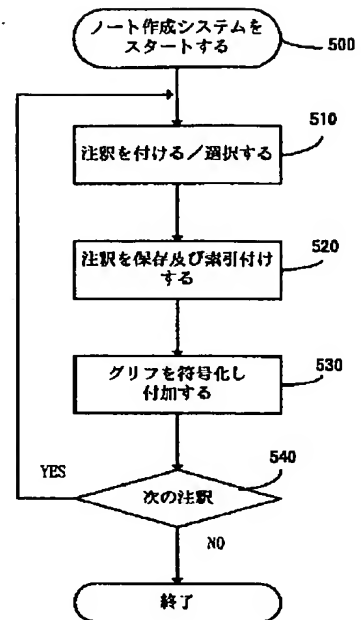
【図2】



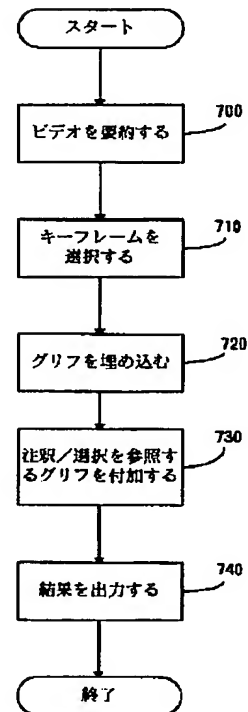
【図3】



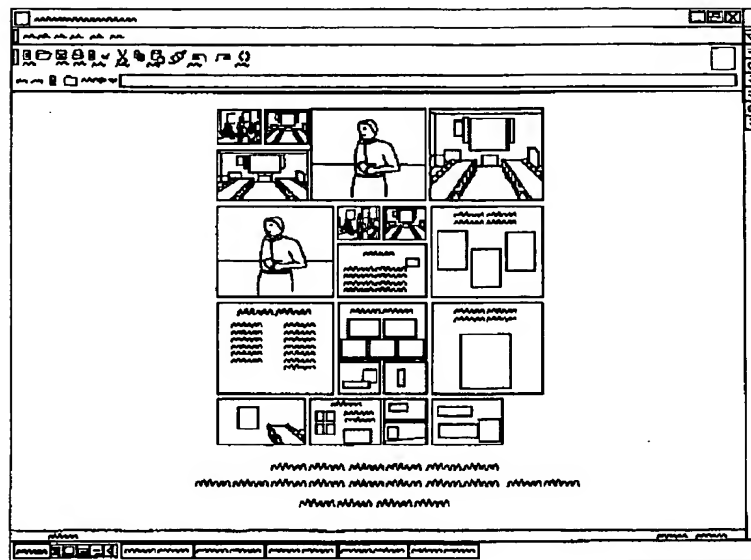
【図6】



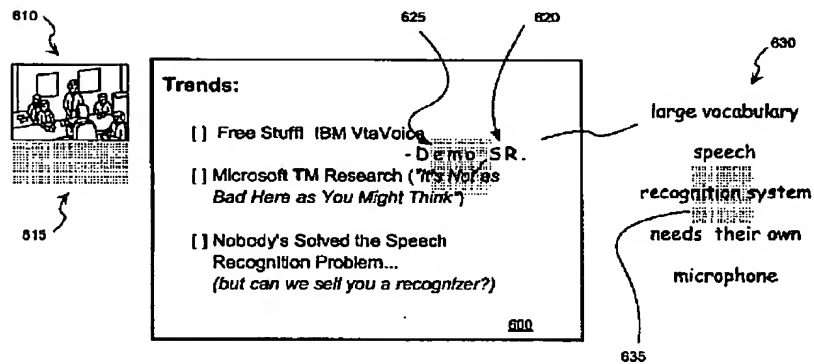
【図8】



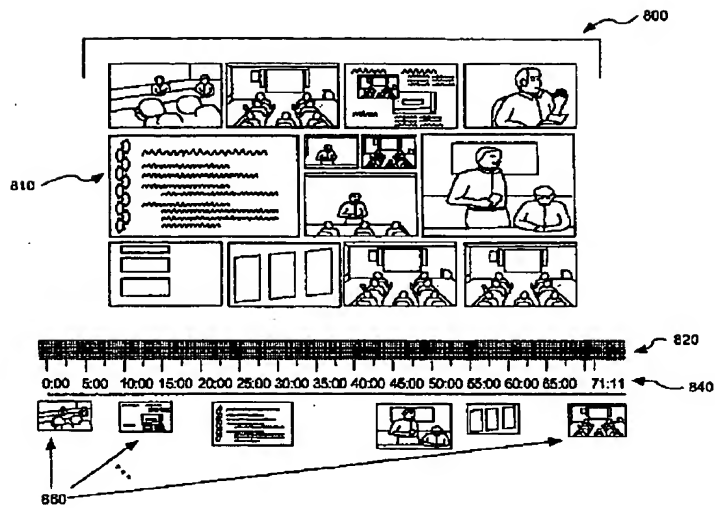
【図4】



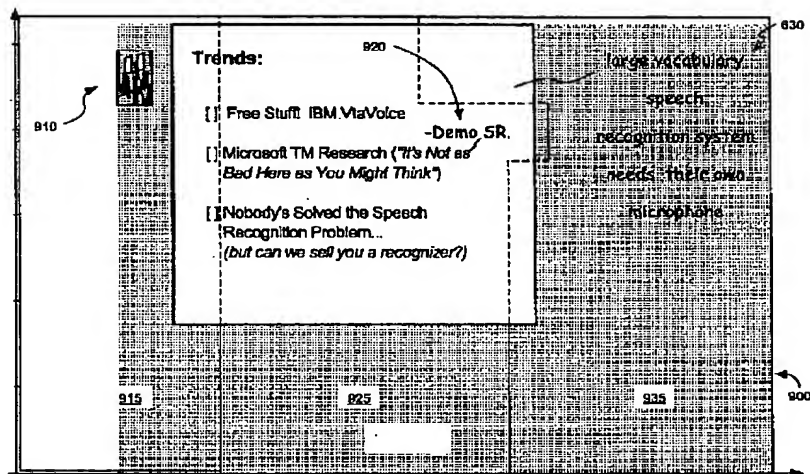
【図7】



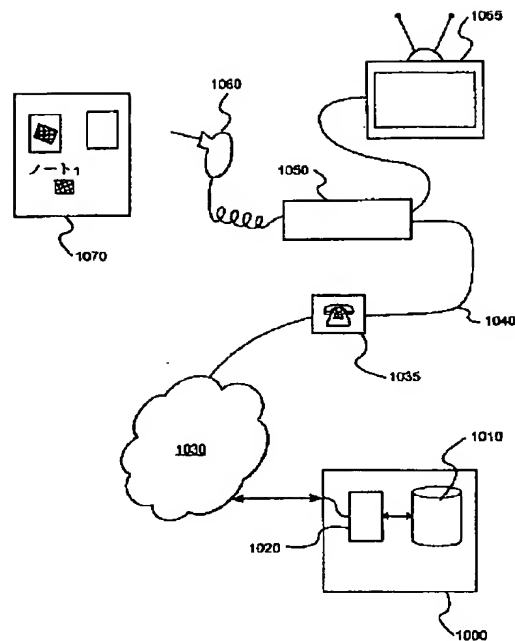
【図9】



【図10】



【図11】



フロントページの続き

(72)発明者 篠岡 信  
日本国 194-0211 東京都 町田市 相  
原町 307-19  
(72)発明者 宮崎 淳  
日本国 251-0027 神奈川県 藤沢市  
鶴沼 桜が岡 3-11-7  
(72)発明者 デイビッド エル. ヘクト  
アメリカ合衆国 94303 カリフォルニア  
州 パロ アルト バーバラ ドライブ  
2001

(72)発明者 エル. ノア フローレス  
アメリカ合衆国 94062 カリフォルニア  
州 ウッドサイド スカイライン ブルバ  
ード 17907  
Fターム(参考) 5C018 AB01 FA04  
5C053 FA14 FA22 FA29 GA11 GB06  
GB21 HA29 HA30 JA21 JA24  
LA06 LA07 LA20  
5D110 AA04 DA02 DA04 DB02 DC01  
DC11 DD01 DE01

【外国語明細書】

SYSTEM AND METHOD FOR VIDEO ACCESS FROM NOTES OR SUMMARIES

Inventors:

Lynn Wilcox  
Patrick Chiu  
Makoto Sasaoka  
Jun Miyazaki  
David L. Hecht  
L. Noah Flores

COPYRIGHT NOTICE

A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

BACKGROUND OF THE INVENTION

Field of the Invention

This invention relates to access and finding a position in video or other recording media. The invention is more particularly related referencing a position within recorded media using a reference marker related to the position. A note, portion of the media itself, or other item related to the position may be utilized as the reference-marker. The invention is still further related to the use of glyph encoding for associating the reference marker to the corresponding

position in the recorded media.

#### Discussion of the background

Multimedia note-taking systems typically capture audio and video during a meeting and slides are created from the captured material. (For example, Tivoli, a system running on LiveBoard, see Moran, T.P., Palen, L., Harrison, S., Chiu, P., Kimber, D., Minneman, S., van Melle, W., and Zellweger, P. "I'll get that off the audio": a case study of salvaging multimedia meeting records. *Proceedings of CHI '97 CM*, New York, pp. 202-209). Tivoli is designed to support working meetings rather than presentation meetings. The ink strokes in Tivoli, which are indexed to the audio, along with any prepared material on the Tivoli slides become the group notes to the meeting. A participant using a laptop may "beam" typed text comments onto a slide in Tivoli.

In a similar example, Classroom 2000, images of presentation slides and audio are captured, but video is not used (see Abowd, G.D., Atkeson, C.G., Brotherton, J., Enqvist, T., Gulley, P., and LeMon, J. Investigating the capture, integration and access problem of ubiquitous computing in an educational setting, *Proceedings of the CHI '98 Conference*, ACM, New York, pp. 440-447; and Abowd, G.D., Atkeson, C.G., Feinstein, A., Hmelo, C., Kooper, R., Long, S., Sawhney, N., and Tani, M., Teaching and learning as multimedia authoring: the classroom 2000 project, *Proceedings of the ACM Multimedia '96 Conference*,

ACM, New York, pp. 187-198). In addition, Classroom 2000 requires effort by the presenter to prepare the slides in a standard graphics format. The slides are displayed on a LiveBoard and note-taking is done with PDA devices pre-loaded with slides. These notes are later synchronized to the audio and the slides which have been annotated by the professor lecturing in front of the LiveBoard.

In yet another example, the Forum (see Isaacs, E.A., Morris, T., and Rodriguez, T.K. A forum for supporting interactive presentations to distributed audiences. *Proceedings of CSCW '94*, ACM, New York, pp. 405-416), is a system uses video as a means for distributed presentations. Everyone, including the speaker, sits in front of a workstation during a presentation. Slides have to be prepared in a specified format. The slides can be annotated with text and marks drawn with a mouse, but the video images cannot be annotated.

In another example, the STREAMS (see Cruz, G., and Hill, R. Capturing and playing multimedia events with STREAMS. *Proceedings of the ACM Multimedia '94 Conference*, ACM, New York, pp. 193-200), is a system for presentation capture that uses video from room cameras. These cameras are also used to capture any presentation content on display. This method has problems when activity in the room obscures the display. Note-taking during the presentation is not supported, although the captured video streams can be annotated during review by adding text comments. None of these systems allow interactive integration

of live images from cameras and presentation material into the notes.

In addition, there are also several known stand alone ink and audio note-taking systems. For example, FXPAL Dynomite (see Wilcox, L. D., Schilit, B. N., and Sawhney, N. Dynomite: A Dynamically Organized Ink and Audio Notebook. *Proceedings of CHI '97*. ACM, New York, pp. 186-193); and Audio Notebook (see Stifelman, L. *The Audio Notebook: Paper and Pen Interaction with Structured Speech*. Ph.D. Thesis. MIT, 1997), which uses paper with audio recording. FiloChat (see Whittaker, S., Hyland, P., and Wiley, M. FiloChat: handwritten notes provide access to recorded conversations. *Proceedings of CHI '94*. ACM, New York, pp. 271-276), is a PC computer with a pen tablet in which audio is indexed with handwritten notes; and NoTime (see Lamming, M., and Newman, W. Activity-based information technology in support of personal memory. Technical Report EPC-1991-103, Rank Xerox, EuroPARC, 1991), was designed to key the user's ink strokes to recorded audio or video.

Also known are video annotation systems. Marquee (see Weber, K., and Poon, A. Marquee: a tool for real-time video logging. *Proceedings of CHI '94*. ACM, New York, pp.58-64) is a pen-based system for making annotations while watching a videotape. A later version of Marquee has modifications to take timestamps on digital video streams from the WhereWereWe multimedia system (see Minneman, S., Harrison, S., Janassen, B., Kurtenbach, G., Moran, T., Smith, I., and van Melle, B. A confederation of tools for capturing and accessing collaborative activity. *Proceedings of the ACM Multimedia '95 Conference*. ACM, New York, pp.523-534). Vanna (see Harrison, B., Baecker, R. M. Designing video annotation and analysis systems, *Graphics Interface '92*. Morgan-Kaufmann, pp. 157-166); and EVA (see MacKay, W. E. EVA: An experimental video annotator for symbolic analysis of video data, *SIGCHI Bulletin*, 21 (2), 68-71. 1989. ACM Press) are text based systems. VideoNoter (Trigg, R. Computer support for transcribing recorded activity, *SIGCHI Bulletin*, 21 (2), 68-71. 1989. ACM Press) displays and synchronizes different streams of activity (video, figures whiteboard drawings, text), but requires post-production to transcribe text from the audio or extract drawings from a whiteboard. These systems are limited by their design based on using videotapes rather than digital video. None of these systems allow interactive integration of video images into the notes. Sharp Zaurus (*Zaurus Operation Manual*. Sharp Corporation, 1996) is a commercial product, which is a PDA with

a digital camera attached. Digital photos can be taken and linked to handwritten notes.

Furthermore, many systems include video playback mechanisms. Fig. 1 illustrates a conventional video playback device 100 that includes a video display area 102. A timeline 105 shows a current position of the video being displayed, and standard VCR controls, including play 115, Fast Rewind 120, Fast Forward 125, Step Back 130, Step Forward 135, etc., are provided that allow a user to navigate the video. However, difficulties arise in finding specific positions within a video, particularly in reference to notes or other items related to the video.

#### SUMMARY OF THE INVENTION

The present inventors have realized that the note-taking systems will have increased ease of playback and greater utilization if important parts of a video or other data stream monitored during notetaking is provided with a simple and accurate retrieval mechanism. Roughly described, the present invention provides a multimedia note-taking system that allows a user to select keyframes or make annotations during a notetaking session (meeting, presentation, or other activity). Each keyframe and/or annotation is captured and indexed to a position of a video of the note taking session or other data stream related to the notetaking. The captured keyframes and annotations are printed on a paper sheet with glyph encoding,

and the printed sheet is used to access the important parts of the video via scanning the glyph codes associated with a target keyframe or annotation and referencing the important part of the video or data stream indexed by the target keyframe or annotation.

These and other objects are accomplished by a system for note-taking with digital video and ink (also referred to as NoteLook). The invention may be embodied in many ways, including, but not limited to any of the following.

An interface for accessing a data stream, comprising, an output, comprising, at least one object that corresponds to at least a portion of said data stream, and a code associated with each object that identifies a portion of said data stream corresponding to the object.

A method of making an index for access to a data stream, comprising, identifying at least one object related to said data stream, determining an offset into said data stream corresponding to each related object, and creating an index that identifies each offset and corresponding related object.

A method of accessing a data stream, comprising, scanning an object corresponding to said data stream, determining an offset into said data stream based on the object scanned, and accessing a portion of the data stream corresponding to the offset determined.

And, a device for accessing video from notes, comprising, a scanning device configured to scan glyph codes embedded in a video index, a computing device configured to read said scanned glyph codes and access a portion of said video based on the glyph codes, and a display mechanism configured to output the retrieved video portion.

The invention also includes a method whereby any of multiple data streams' input into a note-taking device may have portions or frames captured for later reference, and may allow a user to make annotations related to any of the data streams captured (such data streams including any of user inputs including pen-based inputs, a video input, multiple video inputs, data streams, whiteboard inputs, etc.). The method includes making an annotation or selecting a frame from one of the data stream's input, saving and indexing the annotation or selection, and encoding and attaching a glyph that one of identifies the saved selection or annotation.

The steps are repeated for any of multiple selections and/or annotations made during a presentation, or data collection effort. The annotations and selections and attached data glyphs may then be printed out on a paper or placed on other electronic media for selection to initiate a playback of the saved selection or annotation, or begin playing back any of the data streams at a point associated with the annotation or selection and associated glyph. A scanner may be utilized to

scan a data glyph, and data scanned from the glyph is utilized to identify and retrieve the video and an attached (or other) device starts playing the video or other data stream from a location of the data stream identified by the selection and/or annotation associated with the scanned glyph.

The invention may be practiced automatically whereby the data streams are captured and automatically summarized by identifying key frames or other data selections which are then extracted and embedded with glyphs identifying a location in the data stream where the selection was extracted. Glyph referenced annotations and/or selections made by a user manually may also be made separately or included in an output of the selected key frames.

The invention may be practiced over a network, such as the Internet, where a device captures a downloaded or broadcast video stream for annotation or note-taking. Outputs of the invention, such as paper output having a selection and associated glyph, may be scanned by a user, the glyph then identifying an http or other network-related address to a server that then begins streaming the video from the position indicated by the glyph scanned and output on a terminal such as a PC, Palm device, web TV-based, or other display mechanisms.

Various embodiments of outputs of the present invention are described, including block glyphs associated with annotations or images on printed pages, an address carpet of glyphs may be provided in conjunction with an annotated timeline,

and a two-dimensional address carpet with X-Y positions mapped to time and provide selected access to video access similar to a retrieval interface of any device, or a summary interface on a web browser-type product.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings, wherein:

Fig. 1 is a screenshot of a conventional video playback device;

Fig. 2 is a diagram illustrating a note-taking application client running on a pen-based computer;

Fig. 3A is a screenshot of a note-taking-type application having video and whiteboard inputs, and a user annotation area;

Fig. 3B is an example of a Manga video summary according to the present invention;

Fig. 4 is a block diagram of one embodiment of an annotation/data-segment capture device and information flow to produce video summaries according to the present invention;

Fig. 5 is a flowchart that describes a process for capturing and indexing selections and annotations according to the present invention;

Fig. 6 is an example output of glyphed annotations and an

indexed video segment according to the present invention;

Fig. 7 is a flowchart that describes a process of summarizing a video and embedding glyphs according to the present invention;

Fig. 8 is an illustration of a keyframe summary (including video and whiteboard snaps) having an attached timeline and associated data glyph carpet for accessing points in a multi-media data stream;

Fig. 9 is an illustration of an example data carpet mapping multiple points of selected keyframes and annotations captured in a note-taking system to a multimedia data stream; and

Fig. 10 is a block diagram of a hardware configuration being utilized to access points in a video stream over network-connected devices using the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring again to the drawings, wherein like reference numerals designate identical or corresponding parts, and more particularly to Fig. 2 thereof, there is illustrated a pen-based computer 200 running a note-taking application 210. The note-taking application includes a main area 215 resembling a paper notebook page for writing, capturing and, annotating images. There is a small video window 220 for viewing an active video. The active video comes from an active channel, and the user may change channels to view other video streams or present another type of data stream to the user or application. The note-taking application includes facilities to grab a frame that is showing in the video window and place the captured frame (225, for example) in a margin of the main area, or grab a frame as set it as a large background image. "Ink" strokes of a user's pen device 230 are also captured. The thumbnails, background images, and ink strokes are timestamped and are utilized to provide an index into the video or other data streams.

The active video may be provided to the note-taking application via any source. The video source can be captured in a variety of ways: from a room camera or document camera, from a tap into a rear projector, TV, a document camera for overhead slides, VCR or any video stream, or from a small portable camera attached to a pen computer, or any other type of data input device. Multiple video or other data sources may

be provided to the note-taking system.

Fig. 3 illustrates a screenshot 300 of a note-taking system, including a video window 320 and video controls 310. Next to the video window are three buttons for interacting with the video: the top button (channel changer button 322) changes channels, the middle button 324 snaps a thumbnail (thumbnail 370, for example) into the margin of a note page area 315, and the bottom button 326 snaps in a large image 390 (i.e., an image available for annotation or perhaps enlarged for viewing) into the note page area 380.

Further to the left of the video window 320 is a set of VCR-type controls 360. A timeline 350 has a pointer 355 for indicating the current time position of a video being played back. The pointer may also indicate a time position of multiple data streams being captured during a note-taking session (for example, each of a video input, whiteboard input, and pen-based inputs from a user). The timeline can also display indexing information to facilitate browsing and reviewing. The exact configuration of the note-taking device may be of any form and different or additional control mechanisms other than as described herein may be provided.

The system allows users to take notes on the pen-based computer while viewing video that is being digitized and stored on a server (or stored locally or on other devices, e.g. tape, cd, etc.). The notes are time-stamped and synchronized with the video. In addition, the user can snap still images from

the video into the notes, either as thumbnails in the margin or as full-size images in the note page. These images are also time-stamped and synchronized with the video, so that they too serve as indexes.

After a note-taking session, the note-taking system can be used to play back the recorded video. The playback position can be controlled by standard VCR controls such as play forward or reverse. In addition, specific parts of the video can be accessed using the indexes created by the pen strokes and still images in the notes. The user selects the desired index object and presses the PLAY button (or invokes another play mechanism, e.g. Double Click, etc.). An annotated timeline is also available (see Fig. 1) to position the time of video playback.

The above note-taking system roughly describes the NoteLook system developed at PX Palo Alto Laboratory. The NoteLook and other note-taking systems may be utilized to capture images, snaps, annotations, etc., upon which to apply the present invention. Many other types of devices supporting any of capture, annotation, enhancement, or other processes on any of video, audio, multimedia and other data types may also be utilized to apply the present invention. All that is needed is a device capable of capturing any type of data stream and facilities (included in the device or otherwise accessible) for applying the methods and/or processes of the present invention.

It is preferred that frame rate compression be utilized to save bandwidth. Automatic change detection (slide change

detection, for example) may be utilized and only changes in a media stream are sent to the note-taking device. Alternatively, a frame per second rate of material in the media stream may be reduced (i.e., a video shot at 30 fps may be reduced to 1 or 2 fps). This may result in a jittery display at the NoteLook client, however, a perfect video at the note-taking device is not essential to the note-taking process, and, in fact, not needed in situations where the notetaker is present at the video (because the notetaker can hear and see what is taking place). The note-taking device may store the video/data stream or the video/data stream may be captured or maintained on a separate server.

Note-taking systems are particularly useful in a conference room setting, where video from room cameras and projectors can be transmitted to the NoteLook computer over a wired or wireless network. While it is easy to imagine such note-taking devices being available in a conference room, it is less likely that users will have their own devices for use outside the room. Thus it is important to give users an alternate means to access their notes and the recorded video outside the conference room. One of the easiest ways to do this is to provide the user with a printout of the notes. However, this alone does not give access to the stored video.

The present invention embeds data glyphs in any of notes, annotations, clips, snapshots (all collectively referred to as notes), onto printouts or other outputs that provide access to

the video data stream that was utilized during a note-taking session. The glyphs provide time offsets into the video so that a specific section of video can be accessed. In addition, the glyphs can encode a location of the recorded video associated with these notes.

Another application in which glyphs can be used to provide access to video from paper is video summaries. One type of video summary suitable for printing is the Manga-style summary described in Uchihashi et al., U.S. Patent No. \_\_\_\_\_, entitled "AUTOMATIC VIDEO SUMMARIZATION USING A MEASURE OF SHOT IMPORTANCE AND A FRAME-PACKING METHOD," Serial No. 09/276,529, filed March 12, 1999, which is incorporated herein by reference, in its entirety. Here, a video is automatically analyzed and keyframes of varying sized are selected for the summary. Fig. 3B shows an example video summary. The larger keyframes correspond to more important parts of the video. Summaries can be placed on Web pages where they provide access to the video. When a keyframe is selected, the video begins playing from that frame.

In one embodiment, the summaries are printed with embedded glyphs. This allows access to video from paper, rather than through a computer interface on the Web. The user can use a hand scanner to scan a glyph embedded in a selected part of the printed video summary. The scanner is connected to a device such as a Web TV or to a regular PC which then accesses the portion of the video/data stream indexed by the scanned glyph.

The embedded data glyphs in printed notes or in video summaries enable a user interface to the video. Video is accessed by selection of a note or an image captured from the video having an embedded glyph that references a position in the video, a segment of the video, or a set of at least one multimedia or other data streams.

Fig. 4 illustrates a device according to the present invention. A capture device 400 includes at least one input stream of data 410, which may be any of audio, video, whiteboard, or other data inputs. User inputs 420 are utilized to identify portions of any of the data input streams 410 and to make annotations and/or notes associated with any of the input streams or other events of which notes are taken by the user.

The capture device 400 includes a capture/annotation mechanism 425 that captures user-selected portions of the input data streams 410 and any annotations or other notes made by the user, provides those user-selected portions, annotations and/or notes with an ID, index, or timestamp identifying a position in the input data stream, and stores them to a storage device 430. The capture device 400 also includes channels for forwarding the input data streams to the storage device 430 for storage. (Alternatively, input data streams may be provided independently to the storage device 430, or simply maintained on a server.)

A summary mechanism 440 retrieves the captured snapshots, annotations, and notes and formats them in an appropriate format

for output. In addition, the summary device 440 might provide independent summarization and keyframe selection of any one or more of the input data streams 410. The output is formatted for any of a web display such as on a browser device on a personal computer 450, printed out in hard copy form (paper) 455, or stored on an electronic format (floppy disc, for example) 460. The outputs may be provided to any other type of output device including other web-based products, web TV for example, etc.

Fig. 5 illustrates a process according to the present invention. At step 500, a note-taking system is started. The note-taking system includes any one or more of the features as presented above, or is any type of device that receives a data stream that may either be commented, annotated or selected by a user. At step 510, a user of a note-taking system makes an annotation or selects portions of data provided to the note-taking system. At step 520, any annotations or selections made are saved and indexed in a manner that allows retrieval of the selection and/or annotation, and identifies a position in at least one of the data streams provided to the note-taking system. An annotation might also be made in reference to a data stream that is not provided to the note-taking system, but played in conjunction with a note-taking system.

At step 530 a glyph is encoded with information identifying either the annotation or selection and a position in a specific one or multiple data streams, and glyph is attached (or maintained with reference) to the selection or annotation. At

step 540 the annotation/ selection process and glyph attachment is repeated.

As will be appreciated by those skilled in the art, based on the present disclosure, many different variations of the flowchart of Fig. 5 may be implemented to practice the present invention. For example, annotations and selections may be all made and stored in a loop prior to encoding each of those selections and annotations in a batch format. Any of the processes may be performed in parallel or serial depending on architecture and programming techniques. What is important is that annotations and selections are made and glyphs are associated with each of the annotations and/or selections, the glyphs containing information or pointing to information that identifies a portion of a data stream corresponding to the annotations and/or selections.

The glyphs may contain information on the location of the associated video on the server (corresponding to the particular note set), or time offsets into that video, or the glyphs may contain pointers to any other type of picture, video, audio, document, or data. It makes the most sense to use the glyphs to link to a video being displayed at the time a selection, or annotation was made, or to link the glyph to a document being annotated or commented on at that time. Alternately, the user could manually specify the location of the video (via a pop-up window or other mechanism provided by the note-taking device).

Once output onto paper notes, using a scanning device

(hand-held scanner, for example), the user scans image areas in the paper notes, thus designating the desired selection. Glyphs in the scanned image provide appropriate information to initiate playback of the video at the particular time offset.

Various embodiments of glyphs are utilized by the present invention. In one embodiment, a block glyph is provided for each indexed object in a printout or other output. In the example in Fig. 6, there is a glyph 615 associated with the thumbnail (snap) 610, and with each group of ink strokes (ink strokes 620, 630, and respective glyphs 625, 635) on a page of a note-taking device. The objects (thumbnail and ink strokes) were indexed and attached to the respective glyphs according to the process in Fig. 5, or an equivalent. Other types of objects may be similarly indexed and provided with a glyph including any image or data, including background images.

In one embodiment, the location of the video (or other multimedia/data stream) associated with the object is encoded in the attached glyph on the note page. Alternatively, the glyph may provide an entry point into a table that identifies the video and offset as described below.

The ink strokes themselves represent a group of related ink strokes. Each individual pen stroke is not indexed into the video, but instead, the present invention indexes groups of pen strokes. Many methods for grouping pen strokes may be utilized and still fall within the scope of the present invention. Grouping via any of sentence structure (grouping complete

sentences or paragraphs together using character recognition and grammar rules, for example), word recognition (grouping words), or time, for example. However, it is preferred that pen strokes are grouped by time, so that time differences between strokes in the group are small. Thus, an importance of any loss of indexing resolution is reduced. The present inventors have determined that a 1 sec time delta is best for grouping pen strokes.

The design also works for video summaries or any combination of selections, annotations, and video summaries. In the case of video summaries, a block glyph would be associated with each keyframe in the summary.

Fig. 7 is a flowchart of a process of summarizing a video and embedding glyphs according to one embodiment of the present invention. At step 700, a video is summarized. The video may be any video stream input to a note-taking device or separately recorded in conjunction with a conference, meeting, activity, or other event. Video need not be the object upon which the summarization process acts, for example an audio data stream, or other type of product may also be summarized and indexed. For products other than those that are normally represented by a visual image (audio, for example), an icon, graph, or other representation identifying the product may be incorporated.

At step 710, keyframes from the summarized video are selected. Any selection algorithm may be utilized, and/or a set of external criteria may be applied to select the keyframes.

At step 720, the keyframes are indexed into the video or other data stream and associated with a glyph that encodes the indexed position. Other data may also be encoded into the glyph, including an identification of the video or data stream, an http, ftp, or other address where the underlying data stream can be accessed. Alternatively, the glyph may be encoded with a locator number or other identifier that references a table entry where the underlying data stream may be accessed (i.e., step 720 may include building a reference table).

In another alternative, auto indexing may also be utilized to automatically determine important video shots (such as a slide change on a whiteboard data stream) that may automatically be captured and made into specific notes (using change detection criteria, for example). Such auto indexed material may also be linked to a timeline such as shown in Fig. 8, e.g. slide change 810 shown on summary and corresponding timeline snapshot 860 at 25:00.

At step 730, other annotations or selections are also indexed and provided a corresponding glyph. And, at step 740, an output is produced (paper, for example) from which to reference and recall the video and/or data streams indexed by the glyphs.

In another embodiment, the embedded data glyphs are provided in the form of a glyph address carpet. For example, the glyph address carpet may be placed in a horizontal timeline underneath the printed notepage or summary. Fig. 8 illustrates

a video summary 800 with a glyph address carpet 820 and associated timeline 840. The timeline 840 is used to map horizontal position to time, indexed in the video as encoded in a corresponding location of the glyph address carpet 820, linearly or nonlinearly, so that the user can access any point in the video, regardless of whether there is an index object (from a note) or keyframe (from a summary) for that time.

As an enhancement, any objects (from the video summary or notes) may be displayed at appropriate positions on the timeline to aid navigation. Various summary objects (keyframes) 860 are shown in Fig. 8, and are placed in a position relative to an encoded index in the glyph address carpet 820 corresponding to a position of the summary object in the video.

Positioning of the data glyph carpet need only be in conjunction with the timeline to allow a user to perceive a connection between keyframe objects and the time line, such that when an object or place on the timeline is scanned, glyph codes corresponding to the object or place on the timeline are also scanned. Thus, the position of the glyph address carpet 820 may be above the timeline as shown in Fig. 8, placed below the times line, or integrated with the timeline.

Annotations and/or coloring on the timeline or glyph address carpet may be provided to indicate a linear or nonlinear time scale. Multiple parallel address carpet strips can be used to access multiple video (or other data) channels with identical or distinctive (separate or different) time scales. This is

useful when a notetaking device is utilized, such as NoteLook, where multiple video channels may be recorded or snapped from. Multiple strips can also be used to provide extended timescale of a video with spreadout selection.

A spreadout selection allows portions of a video or other data stream to be selected and moved or added to other files. For example, in a video timeline implementation with a glyph data carpet, portions of the timeline (and the embedded glyph) may be selected. These selections may be dragged and dropped into other applications, saved as separate files, played, or any function performed thereon. These functions may be launched by dropping the selection into an icon representing an application configured to perform the function, or via further selection from a pop-up window (e.g. a right-click pop-up window), for example. The function or application would then operate on the video clip represented by the selection.

For the address carpet in the timeline (or other embodiments described herein), a table or other functional mapping may be utilized to map the horizontal position on the address carpet to the time associated with the video. This table could be stored as a glyph code on the paper, which would be scanned and read into a computing mechanism that would then reference any portions of the glyph carpet subsequently scanned to the read table to determine the index into the video corresponding to the portion of the scanned glyph carpet. Alternately, the table could be stored on a server (a server

that also maintains the digitized video or other data, for example). The location of the video can be provided by a separate block glyph.

And, another solution is to integrate a distributed label for the video location into the address carpet in the timeline. For example, both a video name (ID), and a time (position or index) into the video may be encoded into the timeline or other glyph.

In another embodiment, glyph information is embedded in a two-dimensional address carpet. Fig. 9 illustrates one embodiment of a two-dimensional address carpet 900. The glyph address carpet position can be selected to a precision of one glyph, which is typically about a sixtieth of an inch on the page, which is on par with current electronic computer display resolutions. Other resolutions using new or improved glyph type technologies may also be utilized. Although the primary embodiment of the present invention uses a paper interface (printouts) with notes and embedded glyphs, alternate embodiments may utilize any type of output. For example, a computer display having sufficient resolution to display the notes and glyph encodings such that they may be scanned by a user.

In some embodiments, a table is created which maps X-Y locations on the two-dimensional carpet in the neighborhood of note objects (strokes, snaps) to a time in the video. Table 1 illustrates an example table for mapping X-Y positions of note

objects.

X-Y position	data stream ID	Index (timing position)
(1, 4.5)	video1	2252
(6.5, 4)	inks	1.1
(4.25, 4)	inks	1.2

Each X-Y position corresponds to a snap, inkstroke or other object. A selection mechanism for accessing video corresponding to notes or other objects may determine the note or object closest to a position of the scanned 2 dimensional glyph.

Alternatively, the table might be more comprehensive and define a complete range (an area on the 2 dimensional carpet) for accessing the video position associated with an object. For example, video associated with snap 910 might be accessed from from a range of between (0,0) to (2,0) and (0, 5.5) to (2, 5.5) as indicated by box 915. Video associated with ink strokes 920 might be accessed via scanning any glyph within a boundary as indicated by 925, or video associated with ink strokes 930 by scanning any glyph in a boundary indicated by 935. The size and shape of the boundaries may be computed by any method for dividing areas relative to the objects to be accessed.

As in the timeline case, the table correlating glyph codes with indexes into video or other data streams can be stored as a glyph code on the paper, or other configurations, such as stored on a server with the digitized video. In another

alternative, the glyph itself is encoded with all the needed information to access the video.

In yet another table alternative, instead of referencing video streams and index (or offset into the video) by X-Y position, a code embedded in the glyph is read by a scanning device and used to reference the video and any offset or index into the video.

The location of the video can be embedded in the address code as an extended address distributed label or as an interleaved or interlaced data code. For example, data encoded in the glyph such as video ID and position may be encoded in alternating sections of the embedded glyph code. Alternatively, other methods for scrambling or encoding the data in the glyph codes may be utilized, including techniques such as those disclosed in Hecht, U.S. Pat. 5,825,933, entitled "Parallel Propagating Embedded Binary Sequences For Parameterizing Two Dimensional Image Domain Code Patterns in Two Dimensional Space," which is incorporated herein by reference in its entirety. The application of the two-dimensional address carpet may be similarly applied to is similar for video summaries.

Fig. 10 illustrates one embodiment of hardware for applying the present invention in a network environment. A computing device 1000 contains a storage device 1010 that maintains video and/or data streams captured or played during a note-taking session. A server program 1020 is configured to

serve up portions of any of the captured data streams upon request from browser or other types of devices. The computing device 1000 and server program 1020 are each connected to a network 1030 (Internet, for example). A user has a computing mechanism 1050 with attached peripheral scanning mechanism 1060 and display device 1055. The computing mechanism 1050 may be a standard personal computer or a web-TV-based device which is connected (via 1040) to the network 1030. In a typical embodiment, the connection 1040 to the network 1030 is performed through a telecommunications company 1035, and such connection may be any of wired or wireless connections.

An output from the present invention 1070 includes various snaps or notes and corresponding glyphs. A user, utilizing scanning device mechanism 1060, selects and scans a glyph associated with a specific note/snap. Computing mechanism 1050 reads the scanned glyph code and formats an appropriate message that is sent to network 1030 and received by server program 1020. The server program accesses a relevant portion of a data stream stored on storage device 1010 that corresponds to the scanned glyph. The accessed portion is then transmitted via network 1030 back to computing mechanism 1050 and displayed on display device 1055 (computer screen or standard television set).

The computing mechanism 1000, server program 1020 and storage device 1010 may be located on a same computing mechanism as computing mechanism 1000, or may alternatively be maintained on different computing mechanisms also connected via network

1030 or other devices. Connections between any of the servers and computing mechanisms and network 1030 may be any of wired or wireless communications systems, including satellite, microwave, cellular, or other systems.

Computing mechanism 1050 and display device 1055 may take the form as a mobile computing device and display mechanism such as a laptop computer, Palm-type computer or any electronic device that includes a scanner or other mechanism for inputting the glyph codes. The computing device 1050 and display 1055 may be incorporated into a vehicle such as an automobile, delivery or repair van dashboard, or in a hotel lobby or hall adjoining a conference room, or anywhere that the system might be utilized.

In one embodiment, the scanning device mechanism 1060 is integrated into a remote control for a TV and/or Web TV device. Thus, a user may have a printed summary or article, advertisement, or any type of media having embedded glyphs according to the present invention and utilize the scanner to access video or other data associated with the printed summary, article, etc.

The present invention utilizes glyph technology to access video from printed notes and summaries. In addition, this invention includes specific techniques for encoding video location and time offset in glyphs, including table and direct encoding.

The present invention uses glyphs to provide access to video from printouts of notes captured from an electronic

note-taking device (e.g., NoteLook), including:

- associating block glyphs with groups of handwritten strokes and images in a note page;
- using a glyph address carpet to create a timeline for video, thus providing video access from a paper timeline; and
- using a two-dimensional address carpet over the note page, thus providing a paper interface that may be operated by a scanner such as a pen-based scanning device for accessing video from index objects. The paper interface may be produced by printing an output such as Fig. 8 onto a glyph carpet page.

For printed Manga video summaries, as illustrated in Fig.

3B:

- using a glyph address carpet to create a timeline for video, thus providing video access from paper analogous to video access from a timeline on the Web; and
- using a two-dimensional address carpet over the summary page, thus providing a paper interface similar to the Web interface for accessing video from keyframes.

The present invention also provides techniques for specifying the location of the video on the server and the time offset into the video. These allow the video source and time offset to be encoded in the same glyphs, so that separate scans are not required to specify a video source. The present invention includes:

- use of a distributed label in a glyph carpet address to identify source and time position in video; and
- use of interlaced code in the address carpet to identify source and time position.

Experimental results have shown that users may prefer a combination of the timeline and object based access. Further,

the use of paper to access video was well received, and should prove useful for video summaries as well. Overprinting of NoteLook note pages has been done on glyph carpet pages with robust decoding of addresses from handwritten note segments and from video frame prints.

The present invention may be conveniently implemented using a conventional general purpose or a specialized digital computer or microprocessor programmed according to the teachings of the present disclosure, as will be apparent to those skilled in the computer art.

Appropriate software coding can readily be prepared by skilled programmers based on the teachings of the present disclosure, as will be apparent to those skilled in the software art. The invention may also be implemented by the preparation of application specific integrated circuits or by interconnecting an appropriate network of conventional component circuits, as will be readily apparent to those skilled in the art.

The present invention includes a computer program product which is a storage medium (media) having instructions stored thereon/in which can be used to program a computer to perform any of the processes of the present invention. The storage medium can include, but is not limited to, any type of disk including floppy disks, optical discs, DVD, CD-ROMs, microdrive, and magneto-optical disks, ROMs, RAMs, EPROMs, BEPROMs, DRAMs, VRAMs, flash memory devices, magnetic or optical cards,

nanosystems (including molecular memory ICs), or any type of media or device suitable for storing instructions and/or data.

Stored on any one of the computer readable medium (media), the present invention includes software for controlling both the hardware of the general purpose/specialized computer or microprocessor, and for enabling the computer or microprocessor to interact with a human user or other mechanism utilizing the results of the present invention. Such software may include, but is not limited to, device drivers, operating systems, and user applications. Ultimately, such computer readable media further includes software for performing the present invention, as described above.

Included in the programming (software) of the general/specialized computer or microprocessor are software modules for implementing the teachings of the present invention, including, but not limited to capturing and annotating media streams, producing a timeline of significant note-taking events, linking still frames to points in or segments of a media stream and to glyph codes representing the same, recognition of slide changes, summarization of keyframes, producing tables correlating frames, ink strokes, and other data to glyph positions, and communication of results according to the processes of the present invention.

Obviously, numerous modifications and variations of the present invention are possible in light of the above teachings. It is therefore to be understood that within the scope of the

appended claims, the invention may be practiced otherwise than as specifically described herein.

CLAIMS

What is claimed is:

1. An interface for accessing a data stream, comprising:  
an output, comprising,  
at least one object that corresponds to at least a portion of said data stream, and  
a code associated with each object that identifies a portion of said data stream corresponding to the object.
2. The interface according to Claim 1, wherein said codes are embedded in each corresponding object.
3. The interface according to Claim 1, wherein said output further comprises a table that correlates each of said codes to an offset in said data stream.
4. The interface according to Claim 1, wherein said output is a printout.
5. The interface according to Claim 1, wherein said data stream is a video.
6. The interface according to Claim 1, wherein said code

is a data glyph.

7. The interface according to Claim 3, wherein said table is a data glyph.

8. The interface according to Claim 1, wherein:  
each of said codes comprise a data glyph carpet; and  
said objects are placed in positions corresponding to  
their corresponding codes in the data glyph carpet.

9. The interface according to Claim 8, wherein said data glyph carpet is positioned in conjunction with a timeline such that codes in said data glyph carpet correspond to points on said timeline which also represent an approximate offset into said data stream.

10. A method of making an index for access to a data stream, comprising:

identifying at least one object related to said data stream;

determining an offset into said data stream corresponding to each related object; and

creating an index that identifies each offset and corresponding related object.

11. The method according to Claim 10, further comprising

the step of:

outputting each of the identified objects in association with a code that identifies the corresponding offset.

12. The method according to Claim 10, wherein said data stream is a video or multimedia data file.

13. The method according to Claim 10, wherein said object is at least one of a keyframe of said data stream, a selected frame of said data stream, and a user annotation made in reference to said data stream.

14. The method according to Claim 13, further comprising the step of:

grouping sets of related digital ink strokes as a user annotation.

15. The method according to Claim 13, wherein said step of grouping comprises grouping digital ink strokes into sets wherein each digital ink stroke in any one group was created by a user within a predetermined time delta.

16. A method of accessing a data stream, comprising:  
scanning an object corresponding to said data stream;  
determining an offset into said data stream based on the object scanned; and

accessing a portion of the data stream corresponding to the offset determined.

17. The method according to Claim 16, wherein said step of determining comprises:

reading a code associated with the scanned object;  
accessing an offset from a table that correlates the associated code to said offset.

18. The method according to Claim 16, wherein said step of accessing comprises:

sending a request for a portion of said data stream corresponding to said code to a remote computer; and  
receiving said portion of the data stream.

19. The method according to Claim 18, wherein said request comprises one of an http, ftp, or other Internet related request type; and

said remote computer is connected to the Internet and has access to at least said portion of said data stream.

20. The method according to Claim 18, wherein:  
said request includes the code read in association with said object; and

said remote computer accesses a table that correlates said code to the corresponding portion of said data stream.

21. The method according to Claim 16, wherein said step of receiving comprises receiving a streaming video.

22. A device for accessing video from notes, comprising:  
a scanning device configured to scan glyph codes embedded in a video index;

a computing device configured to read said scanned glyph codes and access a portion of said video based on the glyph codes; and

a display mechanism configured to output the retrieved video portion.

23. The device according to Claim 22, wherein:

said computing device include a network connection capable of connecting to any of a wired or wireless network; and

said computing mechanism is further configured to request said portion of said video from a remote device via said network connection.

24. The device according to Claim 23, wherein said computing mechanism includes a table that correlates said glyph codes to an offset into said video.

25. The device according to Claim 24, wherein said

computing mechanism is further configured to read said table via said scanning device.

26. The device according to Claim 22, wherein said computing device comprises a palm type computer and a wireless network connection.

27. The device according to Claim 22, wherein said scanning device is integrated into a remote control for auxiliary equipment.

28. The device according to Claim 27, wherein said auxiliary equipment is one of a computer, a television, and a Web-TV device.

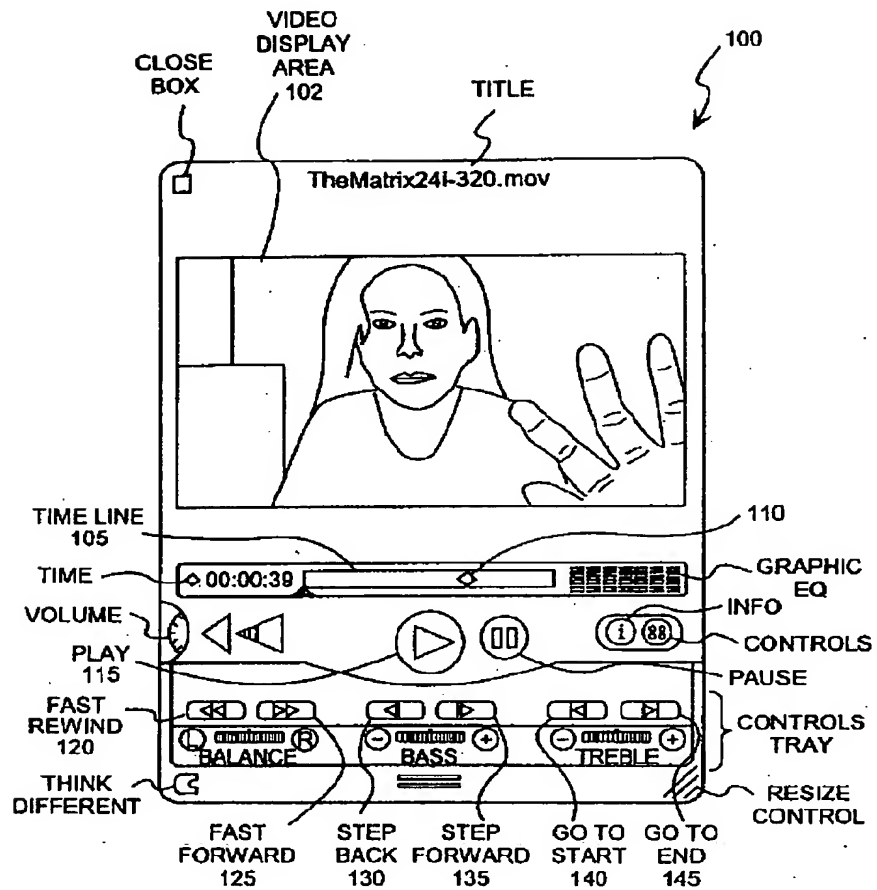


FIG. 1

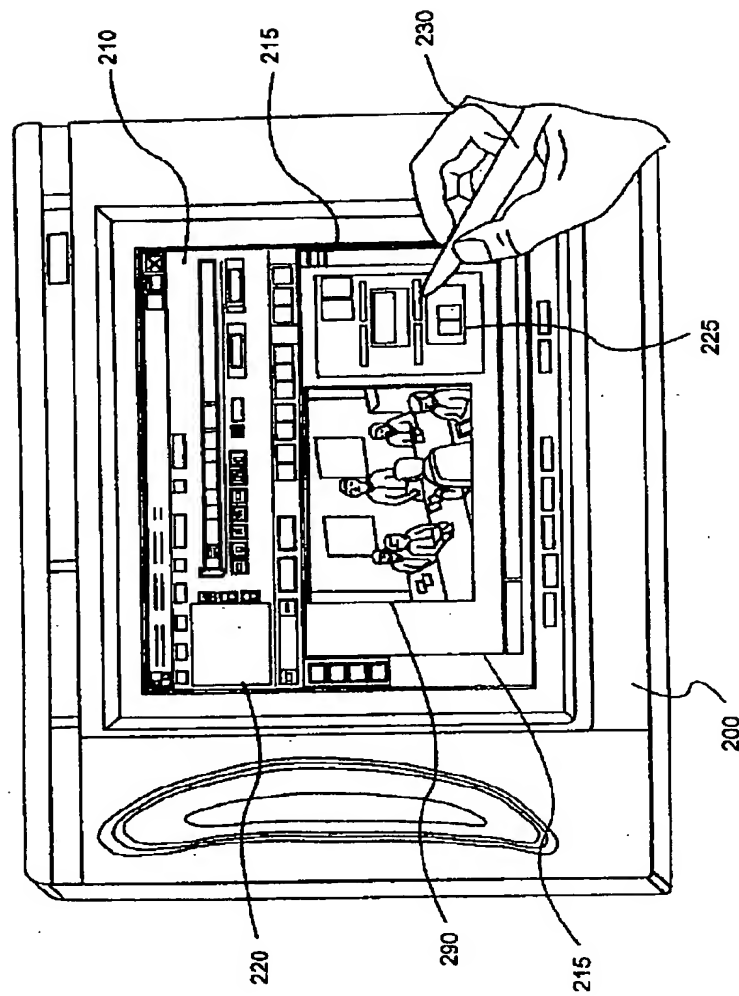


FIG. 2

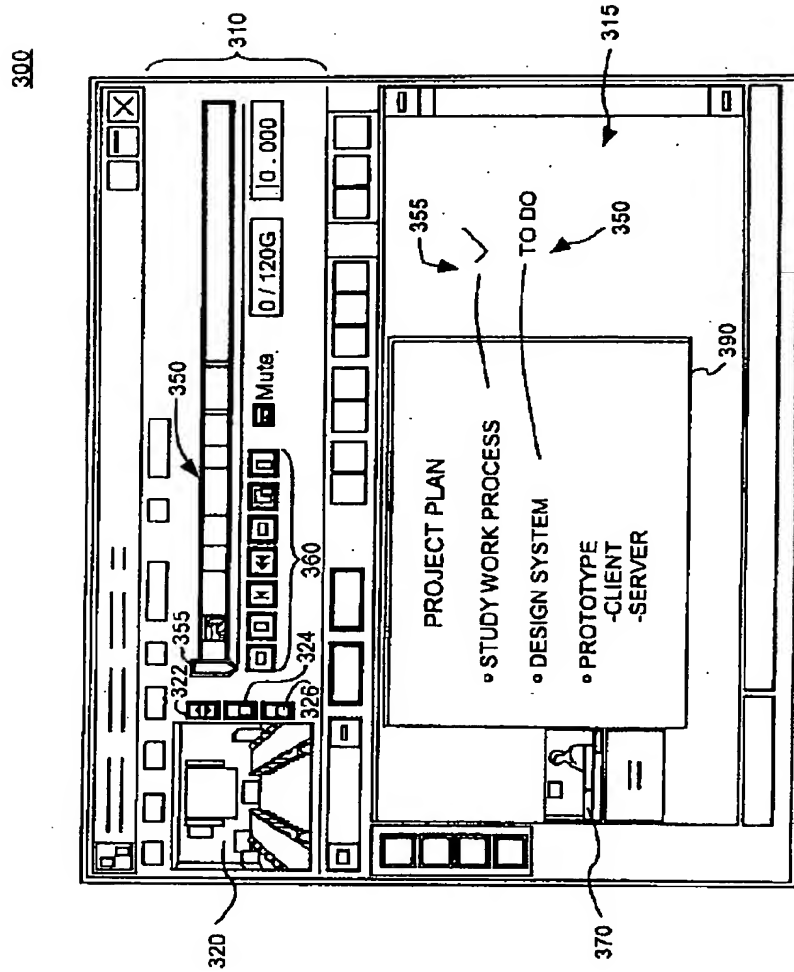


FIG. 3A

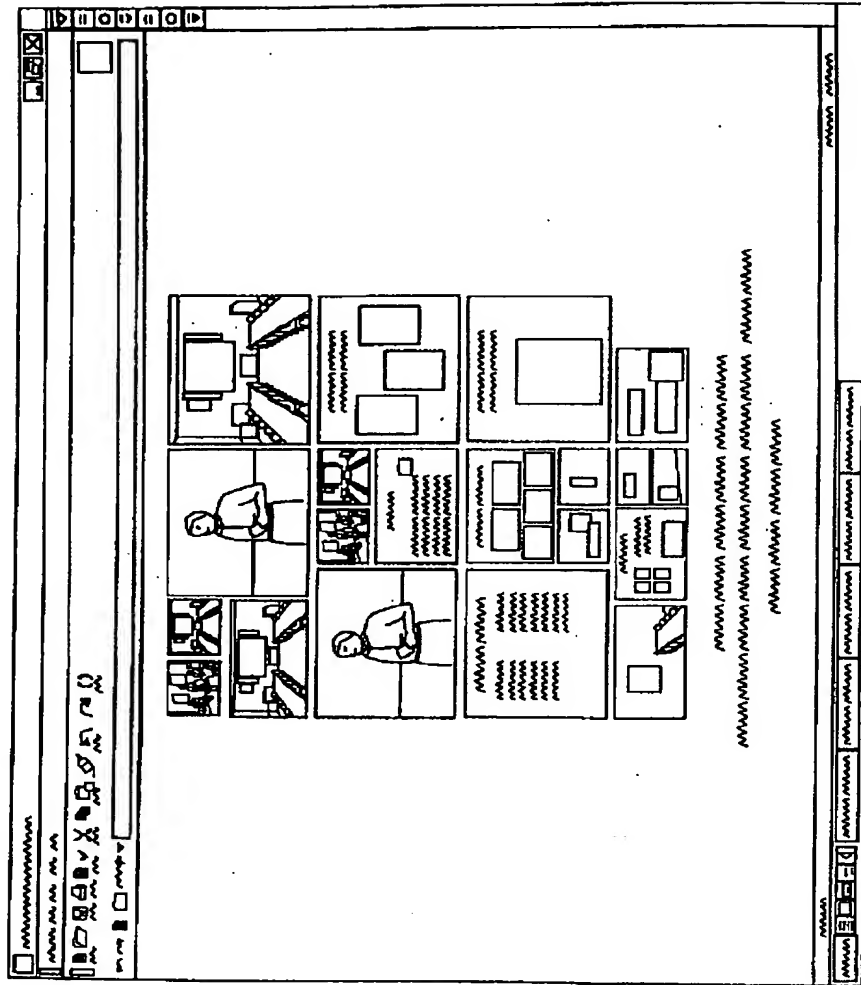


FIG. 3B

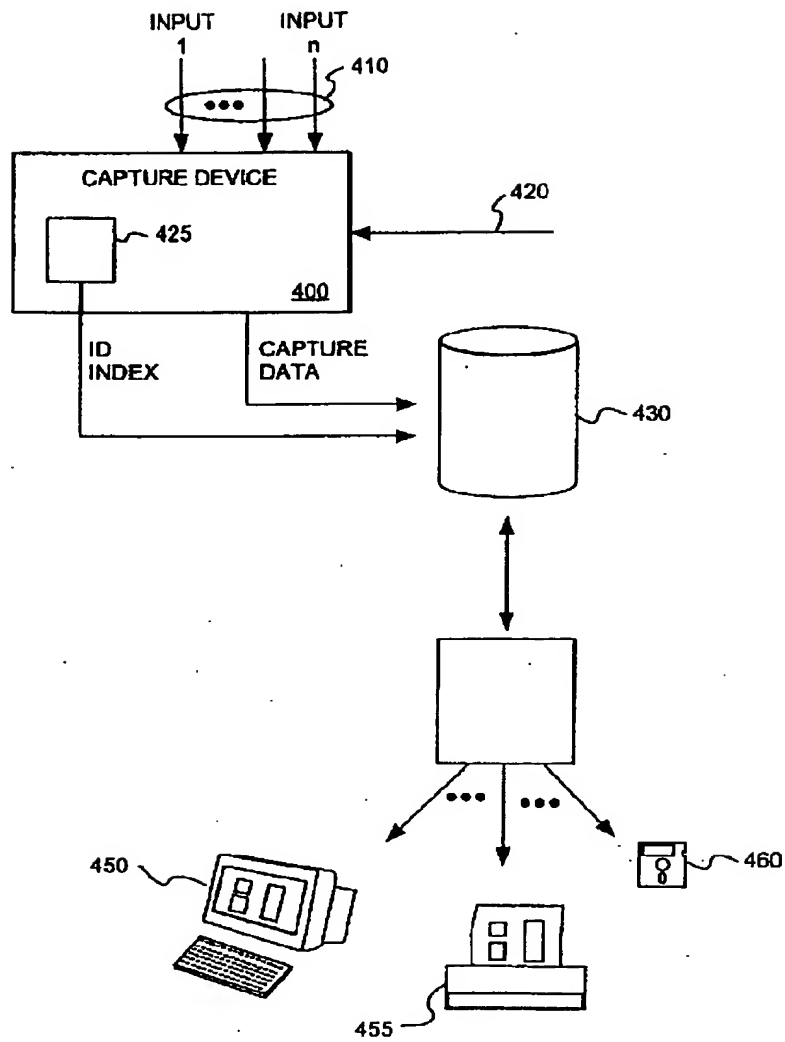


FIG. 4

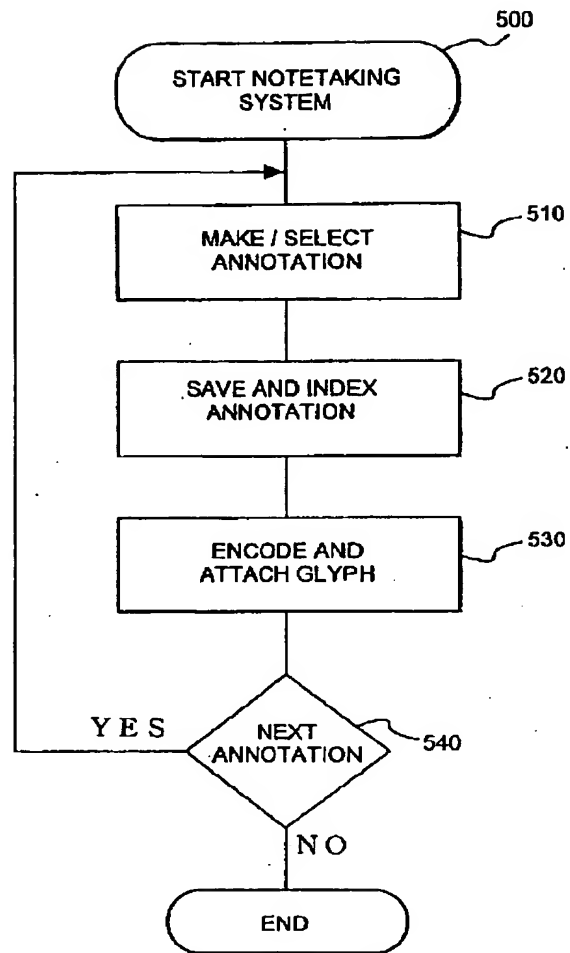


FIG. 5

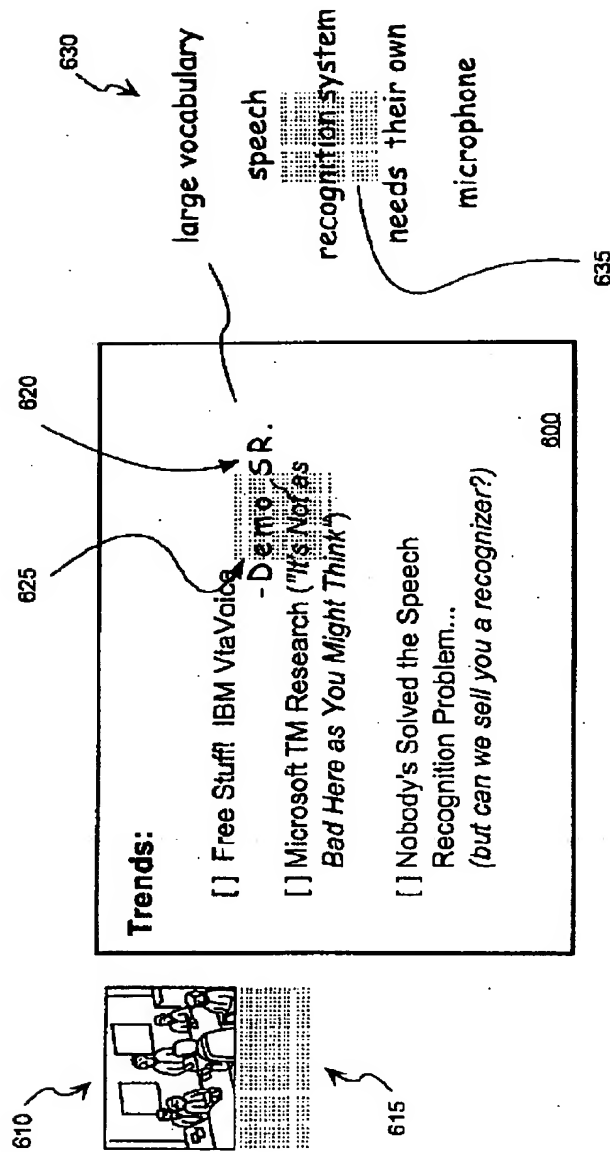


FIG. 6

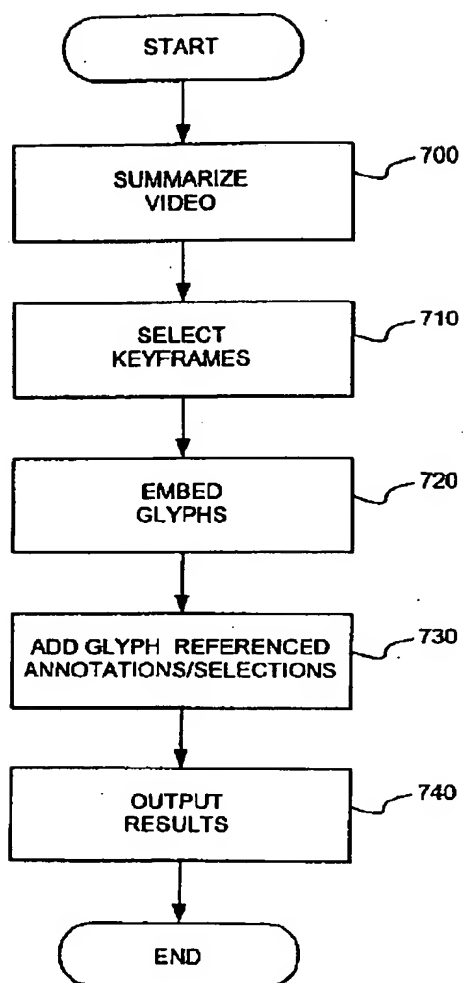


FIG. 7

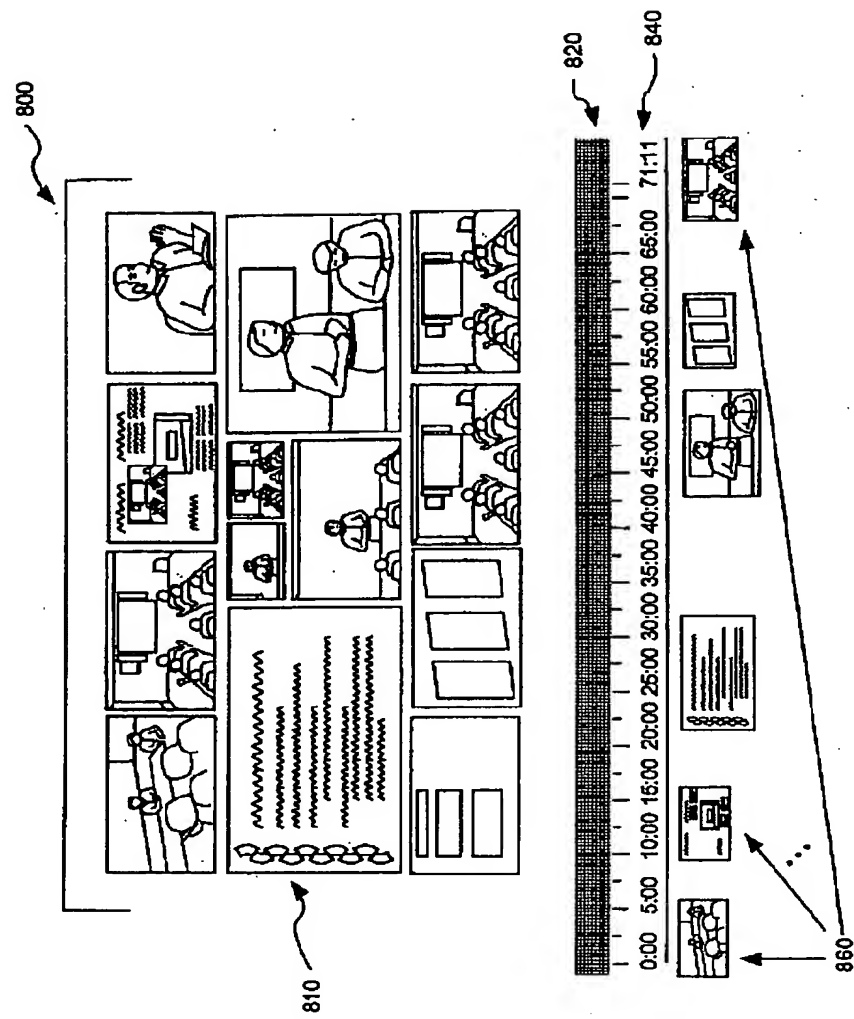
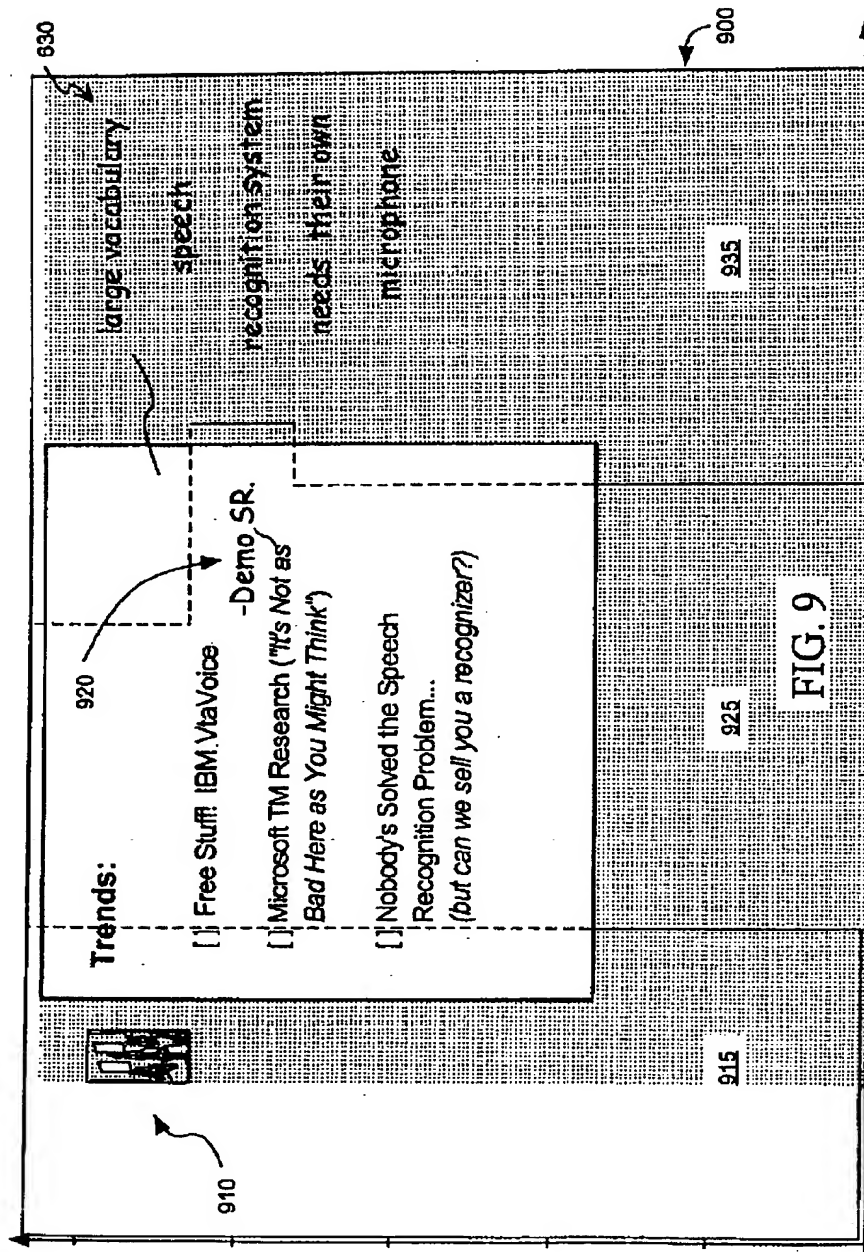


FIG. 8



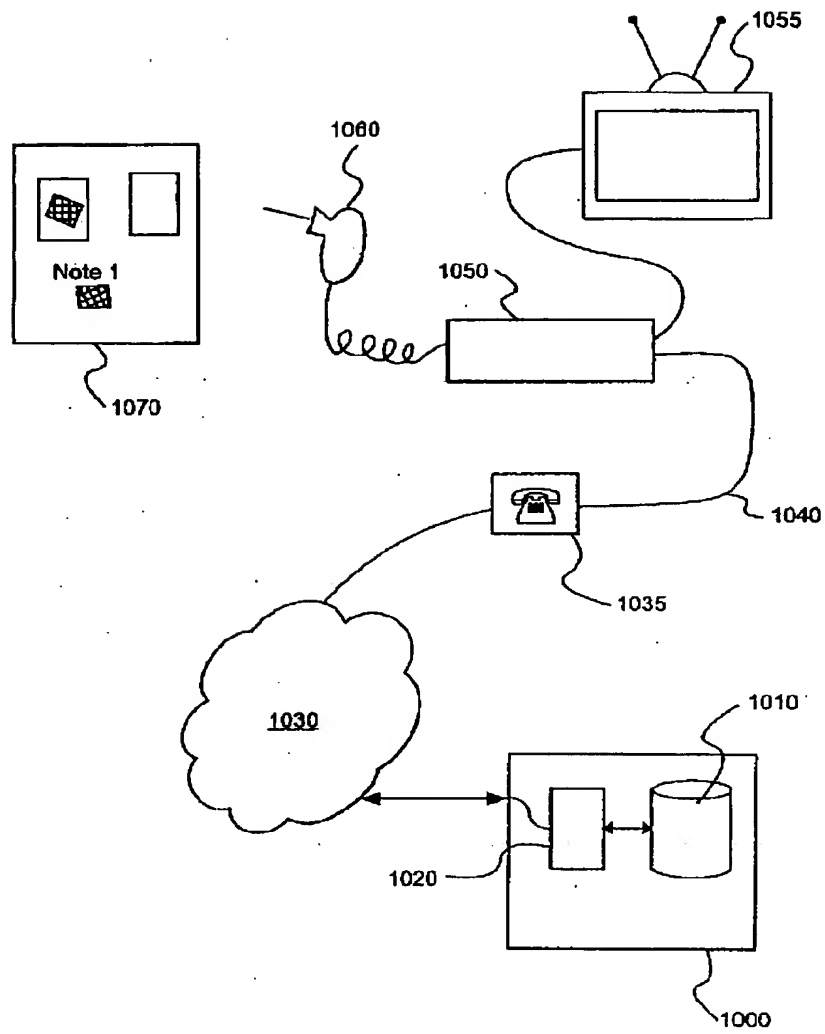


FIG. 10

**ABSTRACT**

Recorded video is accessed from printed notes or summaries derived from the video. Summaries may be created automatically by analyzing the recorded video, and annotations are made by a user on a device for note-taking with digital ink and video. The notes and/or summaries are printed along with data glyphs that provide time based indexes or offsets into the recorded video. The indexes or offsets are retrieved by scanning the glyph on the printout. The glyph information can be embedded in the printouts in many ways. One method is to associate block glyphs with annotations or images on the printed pages. Another method is to provide an address carpet in an annotated timeline. Yet another method is to provide a two-dimensional address carpet with X-Y position mapped to time which can be used to provide selected access to the video. The accessed video may be played back on the note-taking device on a pen computer, or on a summary interface on a Web browser-type device.

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-057981

(43)Date of publication of application : 22.02.2002

(51)Int.Cl.

H04N 5/91  
G11B 27/00  
H04N 5/7826  
H04N 5/92

(21)Application number : 2001-115802

(71)Applicant : FUJI XEROX CO LTD

(22)Date of filing : 13.04.2001

(72)Inventor : WILCOX LYNN D  
CHIU PATRICK  
SHINOOKA MAKOTO  
MIYAZAKI ATSUSHI  
HECHT DAVID L  
FLORES L NOAH

(30)Priority

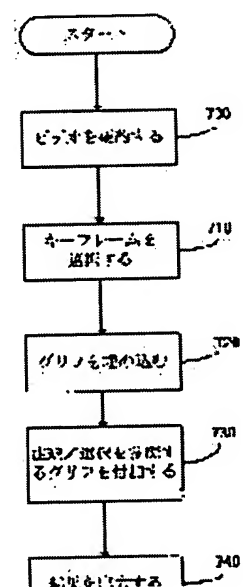
Priority number : 2000 548685    Priority date : 13.04.2000    Priority country : US

(54) INTERFACE TO ACCESS DATA STREAM, GENERATING METHOD FOR RETRIEVAL FOR ACCESS TO DATA STREAM, DATA STREAM ACCESS METHOD AND DEVICE TO ACCESS VIDEO FROM NOTE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a video monitored during note generation or an important part of other data stream with a simple and accurate retrieval mechanism.

SOLUTION: A recorded video is accessed from a printed note or summary introduced from the video. Analyzing the recorded video automatically generates a summary and a user for a device generating a note in digital ink and video provides a remark. The note and the summary are printed together with a data glyph providing a time-



based index or offset to the recorded video. The index or offset is retrieved by scanning the glyph of printout. Glyph information can be imbedded in the printout by many methods. The accessed video is reproduced by a summary interface of a note generating device or a device of a web browser type.

---

## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office